PSYCHOLOGY IN RUSSIA: STATE OF THE ART

Volume 19 Issue 3 2026

Clinical psychology

Age-Related Trajectories in Symptom Manifestation in Preschoolers with Autism and Developmental Delay 3
Nasledov, A.D., Tkacheva, L.O., Miroshnikov, S.A.

Organizational psychology

Working from Home Across National Contexts: Demands, Resources, Engagement, and Productivity Among Italian and Russian Employees 25
Toscano, F., Zappalà, S., Polevaya, M.V., Kamneva, E.V.

Developmental psychology

Teacher Implementation of Sociodramatic Play in Early Childhood Education: A Systematic Review 45
Sukhikh, V.L., Veresov, N.N., Gavrilova, M.N.

Cognitive psychology

Experimental Modelling of Body Schema with the Rubber Hand Illusion and the Hand Laterality Judgment Task: Cognitive and Meta-cognitive Levels 73
Dobrushina, O.R., Arina, G.A., Kazantseva, D.A., Tsypushtanova, M.M., Makarova, A.G., Dobrynina, L.A.

Educational psychology

Self-Monitoring Skills in Online Learning and the Role of Metacognitive Prompts: An Exploratory Study 89
Koval, A.A., Gorbunova, A.Y., Podolsky, A.I., Getman, A.V.

Psychometrics

Psychometric Characteristics of the Russian-Language Version of the Posttraumatic Growth and Depreciation Inventory — Expanded Version (PTGDI-X) 106
Shishkova, A.M., Bocharov, V.V., Valieva, T.V., Tedeschi, R.G.

Personality psychology

The Values System of Russian University Students Compared with Their Level of Aggression 131
Konovalov, I.A., Linkov, A.L., Rean, A.A.

Psychology in Russia: State of the Art
Volume 19, Issue 3, 2026

Russian Psychological Society
Lomonosov Moscow State University

CLINICAL PSYCHOLOGY

Age-Related Trajectories in Symptom Manifestation in Preschoolers with Autism and Developmental Delay

Andrey D. Nasledov[a], Liubov O. Tkacheva[a*], Sergey A. Miroshnikov[a]

[a]*St. Petersburg State University, Russia*

*Corresponding author. E-mail: l.tkachewa@spbu.ru

Background. Symptoms of autism spectrum disorder (ASD) and developmental delay (DD) often overlap in preschool children, yet their manifestation may change with age. While previous studies have described symptom profiles in separate age groups, little is known about how these symptoms unfold longitudinally across the preschool period. In particular, comparative trajectories of symptom manifestation between children with ASD and those with DD across three distinct age ranges remain underexplored. Understanding such trajectories is critical for early differential diagnosis and intervention planning.

Objective. To understand age-related changes in symptom presentation, which is essential for differentiating between ASD and DD, as well as for selecting optimal intervention strategies.

Design. A three-factor model of autism symptoms was established for 1–2-year-old children to capture age-specific patterns of symptom manifestation, which were compared with those observed in 3–5- and 6–7-year-olds with ASD. The overall structure of the model and its measurement equivalence were confirmed across the full preschool age range. Factor scores were calculated as the mean of consistent items for each participant. Children with ASD and DD were then compared on these factor scores to examine age- and diagnosis-related trajectories in symptom expression.

Results. The three symptom domains — Communication Disorders, Sensory Disintegration, and Hyperactivity — persisted throughout the preschool years in children with ASD, with Communication Disorders showing the highest prevalence in the youngest age group (1–2-year-olds). Hyperactivity and Sensory Disintegration were commonly observed in both children with ASD and those with DD at age 1–2 years. However, while all three symptom domain decreased across preschool age in children with DD, they remained pronounced in children with ASD.

Keywords: autism, developmental delay, preschoolers, age-related changes, developmental trajectories

© Nasledov, A.D., Tkacheva, L.O., Miroshnikov, S.A., 2026

ISSN 2074-6857 (Print) / ISSN 2307-2202 (Online)
http://psychologyinrussia.com

Conclusion. The observed patterns of symptom manifestation reflect fundamental and stable neurodevelopmental differences, shaped by a range of interacting factors. These patterns also capture developmental shifts, in which the prominence of specific symptom domains changes with age and experience. Among the symptom domains examined, Communication Disorders emerged as the most distinguishing feature for differentiating children with ASD from those with DD in early childhood.

Introduction

Given the high prevalence of autism (Isaac et al., 2025), understanding the age-related manifestation of autism symptoms in early childhood has become a priority in contemporary research for several compelling reasons. First, early childhood is a critical window of opportunity due to significant neurodevelopmental plasticity (Mualem et al., 2024). Identifying autism symptoms during this period enables timely interventions, thereby maximizing developmental outcomes, as demonstrated in multiple studies (Maksimović et al., 2023). Second, specific early intervention, such as applied behavior analysis (ABA) and speech therapy, have been shown to significantly improve cognitive, social, and communication skills in children with autism spectrum disorder (ASD) (Gitimoghaddam et al., 2022). Characterizing how symptoms manifest across different ages helps refine diagnostic criteria and facilitate the identification of children most likely to benefit from such interventions. Third, early identification and intervention may prevent or mitigate the emergence of secondary emotional and behavioral difficulties that often arise when autism goes unrecognized or inadequately supported (Esposito et al., 2024).

Another motivation for this study stems from the well-recognized heterogeneity of ASD. It is widely accepted that autism is a highly heterogeneous condition, and this diversity is particularly pronounced in early childhood (Masi et al., 2017). Understanding how symptoms evolve with age can therefore inform the identification of subtypes and the tailoring of interventions to individual profiles. Research into age-related symptom dynamics is essential, as the severity of autistic symptoms is known to change over time (Waizbard-Bartov & Miller, 2023). Tracking these patterns allows researchers to examine how specific symptom domains, such as repetitive behaviors or social communication difficulties, manifest across the child's development, particularly as environmental and social demands shift with age. Furthermore, characterizing symptom manifestation in early childhood may enable the detection of subtle, earlier markers of autism, thereby supporting earlier and more precise diagnosis. Another benefit of this line of research lies in its potential to refine diagnostic criteria for ASD. By illuminating how symptoms manifest across different developmental stages, such work can make diagnostic criteria more accurate and sensitive to the developmental nuances of the condition.

The findings of this study address age-specific challenges in assessment. Diagnostic tools and procedures can be adapted to better suit the unique context of evaluating very young children. This, in turn, may facilitate the identification of age-appropriate assessment methods, while minimizing false positives, false negatives, and the high

variability in sensitivity and specificity that often accompany traditional instruments. Such variability is known to arise from multiple factors, including floor effects, which have been documented in widely used tools such as the Modified Checklist for Autism in Toddlers, Revised with Follow-Up M-CHAT-R/F; Communication and Symbolic Behavior Scales Developmental Profile CSBS DP; Autism Diagnostic Observation Schedule, Second Edition ADOS-2 (Fulceri et al., 2025; Saldaris et al., 2024; Thurman et al., 2024; Zheng et al., 2024). Understanding how autism symptoms manifest at different ages is also crucial for differentiating ASD from other developmental conditions that may present with overlapping features in early childhood. These include speech and language delays, intellectual disabilities, global developmental delay, attention-deficit/hyperactivity disorder (ADHD), sensory processing disorder and separation anxiety. The importance of such differentiation is underscored by the high rates of comorbidity between ASD and other conditions in early childhood (Al-Beltagi, 2021).

The objectives of this study were twofold. First, we aimed to identify age-specific patterns of autism symptom manifestation in younger (1–2 years), middle (3–5 years), and older (6–7 years) preschoolers with ASD. Second, we sought to compare these age-related trajectories with those observed in preschoolers with DD. By doing so, this study intended to support early identification and intervention, deepen understanding of developmental trajectories in ASD, inform diagnostic practices, and optimize intervention strategies, ultimately contributing to improved outcomes for individuals with autism and their families.

Methods

Participants

Data from the youngest age group (1–2 years) were combined with data from the middle (3–5 years) and oldest (6–7 years) age groups. *Table 1* presents the sample composition by age, sex, and diagnosis.

Table 1

Composition of the Combined Sample

		Diagnosis		Total
		ASD	DD	
Age	1–2-year-olds	130 (89 male)	27 (22 male)	157 (111 male)
	3–5-year-olds	487 (345 male)	218 (148 male)	705 (493 male)
	6–7-year-olds	300 (217 male)	77 (44 male)	377 (261 male)
Total		917 (651 male)	322 (214 male)	1,239 (865 male)

Sample size was determined followed the recommendations of S.B. Green (1991). Given that confirmatory factor analysis (CFA) was planned with no more than 10 observed variables per factor, applying the rule of at least 10 observations per independent variable yielded a minimum required sample size of 100. Applying Green's

alternative rule, based on a power of 0.80 (α = .05) and up to 10 predictors per factor, a minimum of 119 observations was required to detect an average effect size. Consistent with this assumption, factor loadings exceeded |.5|, suggesting an effect size at or above average. Thus, the sample of 1–2-year-olds with ASD (N = 130) is sufficient for the study's objectives.

Materials

For this stage of the study, we developed an online questionnaire assessing autism markers in 1–2-year-olds, modeled after instruments previously developed for 3–4- and 5–6-year-old children (Nasledov et al., 2021; Nasledov et al., 2024a). The questionnaire comprises 274 markers of ASD, organized into tasks based on the situational contexts in which these markers typically manifest (*e.g.*, play, dressing, communication). Each marker (item) is rated on a binary scale (0 = no — marker not present, 1 = yes — marker present). Consistent with our previous approach, the item bank represents a deliberately extensive collection of descriptions of autism signs and their situational contexts. Items were derived from survey responses of professionals working in medical, psychological, and educational commissions, who reported the markers they rely on during diagnostic assessments. The interface of the questionnaire mirrors the versions used for 3–4- and 5–7-year-old children, but includes only markers applicable for the assessment of 1–2-year-olds. Alongside new tasks developed specifically for this youngest age group, the questionnaire incorporates 36 items that were also included in the instruments for 3–7-year-olds. Of the 40 ASD screening items originally used for 3–4-year-old children (Nasledov et al., 2021; *Table 4*), 36 were retained for analysis in this study. Four items were excluded from the current analysis due to their inapplicability to 1–2-year-old children: B2501, C3304, L1304, and S5301.

Procedure

The data collection procedure was consistent with that previously employed for 3–4- and 5–6-year-old children (Nasledov et al., 2021; Nasledov et al., 2024a). Data were collected between November 2024 and March 2025 by 23 experienced specialists (psychologists, speech pathologists), who worked regularly with the children. Data collection was conducted via an online questionnaire and did not require tests or other assessments beyond routine inquiries. Specialists were asked to respond to items based on their routine observations and checkups regarding autism markers detectable in each child. Parents were interviewed through face-to-face conversations. To ensure objectivity and minimize bias, specialists completed the questionnaire only for children with whom they had already held multiple meetings and parent consultations — typically at least five sessions. The team of specialists consisted of professionals who had participated in previous studies.

Group assignment (ASD vs. DD) was determined by the same specialists working directly with the children, based on their prior experience and current expert opinion derived from practical work with each child. Parents of 1–2-year-olds sought correctional support rather than a formal diagnosis, which is typical for this age group. At approximately two years of age, children undergo rapid development and

exhibit a wide variation in behavior, communication, and social skills. Some typical developmental delays or differences may resemble early signs of autism, but may resolve naturally or follow different developmental trajectories. Consequently, distinguishing between typical variation and a neurodevelopmental disorder at this age is challenging, particularly when a child is assessed by unfamiliar individuals in novel situations, which may elicit uncharacteristic behaviors. Thus, data were collected on children without a pre-existing formal diagnosis. Children were classified by specialists into ASD or DD groups based on observations during scheduled classes and consultations, conducted at the request and with the consent of parents. A detailed description of the online survey and parent interview procedures is provided in previous articles (Nasledov et al., 2021; Nasledov et al., 2023; Nasledov et al., 2024a). In total, 157 1–2-year-old children participated in the study (130 with ASD, 27 with DD).

Data Analysis

The data analysis was carried out in two stages: first, a factor structure with factor validity and invariance in different parts of the sample was identified, and then the calculated factors were used to identify the features of ASD symptoms in children of different age groups, and depending on the diagnosis (ASD or DD). At the first stage, exploratory factor analysis (EFA) was used first, followed by confirmatory factor analysis (CFA), including intergroup analysis. EFA was used on tetrachoric correlations using the minimum remainder method, with parallel analysis to estimate the number of factors. The "diagonally weighted least squares (DWLS) method using a polychoric correlation matrix" was used for CFA, which does not require multidimensional normality (Li, 2016). Confirmatory factor models were evaluated using recommended consent indices and their thresholds (Byrne, 2010; Chen, 2007; Kline, 2011). At the second stage, the obtained factors were calculated as the average values of the items included in them, and samples differing in age and diagnosis were compared according to the calculated factors. To analyze the differences in symptoms depending on age (1–2, 3–5, and 6–7 years), One-Way Multivariate ANOVA was used: factor – Age; dependent variables – calculated factors. To analyze the differences in age-related changes depending on the diagnosis (ASD, DD), Two Way Multivariate ANOVA was used: factors – Age, Diagnosis; dependent variables – calculated factors. In both cases, Post Hoc Comparisons and Contrasts within the framework of one-dimensional ANOVA were used to clarify the statistically significant multidimensional effects of factors and their interactions. The entire analysis was performed using JASP 0.18.3 and Jamovi 2.4.14.0 software.

Results

Exploratory factor analysis (EFA) was performed on the full sample of children with ASD ($N = 917$) using 36 binary items to identify the underlying factor structure and group items into subscales. The analysis employed the minimum residual extraction method, a tetrachoric correlation matrix and oblimin rotation. Following the removal of 10 items and reduction to three factors, a parsimonious three-factor structure was obtained comprising 26 items (8–9 per factor). Each item loaded at ≥ .531 on its

primary factor and ≤ .301 on all other factors. Each factor demonstrated adequate internal consistency, with Cronbach's alpha coefficients ranging from .781 to .863. The factor composition is shown in *Table 2*.

Table 2

Main Results of Factor Analysis and Verification of Reliability of Scales

Factorization of 26 items ($N = 917$): $\alpha = .891$; 59.3% of Variance	FL*
Factor 1: "**Communication**" (*Com***; 28.4 % of Variance), $\alpha = .863$ (9 items)	
L8201. The child is practically unable to communicate	.988
M7205**. The child easily talks about their needs. Uses words, phrases, gestures, and facial expressions.	.855
L8303. The child does not respond to questions addressed to him/her.	.847
M1601. The child does not attempt to communicate with other people. He/she does not communicate his/her needs, tries to take everything on his/her own, or uses strategies typical to young children (crying). What the child was "asking" for becomes clear when the crying stops.	.753
G2101. The child plays by him/herself, is aloof, does not allow others to join his play, neither adults nor children.	.747
L0903**. The child immediately looks at the person who is addressing him/her.	.741
L1003**. Showing something to another person, the child places the object so that it can be viewed and checks whether the person sees what he/she is being shown.	.731
L1501. The child does not imitate the actions of other people: he/she is busy with his/her own business, and does not pay attention to others.	.715
C3201. It seems that the child does not pay attention to the surroundings, but wanders around, takes objects aimlessly, does not focus on them, and immediately throws them, sometimes behind his back.	.677
Factor 2: "Hyperactivity" (*Hyp****; 17.8% of Variance), $\alpha = .795$ (8 items)	
N6901. There are too many unnecessary movements in the child's activity; the child is fussy, hyperactive.	.798
B2502. Can't play quietly.	.787
B2901. The child is aggressive, pugnacious, and prone to physical aggression against animals and other people.	.743
B6202. When something blocks the child from getting what they want, the child "loses control" (*e.g.*, meltdown, aggression, prolonged crying) beyond what is expected for the child's age.	.740
P3905. The child can't sit still, leaves his seat in the classroom or elsewhere, jumps up, and wanders around.	.731
B6201. Loses self-control, is prone to emotional outbursts.	.718
B2603. The child is often angry and irritable.	.693
B2503. The child is challenging to control. His/her behavior depends on external stimuli - "I run everywhere I look" (chaotic behavior).	.598

Factor 3: "**Sensory Disintegration**" (*SD****; 16.6% of Variance), α = .782; (9 items)	
B2804. The child prefers rituals (goes to bed under a certain scenario, drinks only from his/her favorite cup, dresses only in a specific order, walks only along a certain route).	.763
B6501. The child is overly attached to certain objects (blanket, toy, and clothing). If a favorite item is lost, he/she is worried, may get hysterical.	.759
F2303. The child is rebellious, demanding to get changed into the clothes that he/she prefers.	.736
B2805. The child is very picky about food, has a limited range of favorite dishes, constantly demands to be given the same food, and has requirements for food shape, color, consistency, design.	.615
N4205. The child likes to watch lights turn on and off, doors open and close, wheels turn, fan spinning, blinds open and close, shiny objects, pages flickering when flipping through, etc.	.594
I0103. The child retains an unusually long interest in certain objects toys, ropes, balls, stones, plugs and lids, car wheels, toy parts, etc.	.575
N4601. The child is picky/sensitive to certain food textures (for example, pieces in mashed potatoes or porridge). Such foods are perceived by the child as disgusting or intolerable, leading to active avoidance.	.575
N4602. The child's mouth and the space around the mouth are hypersensitive (difficulties in teeth brushing, speech therapy massage, visiting a dentist).	.534
B2707. The child has unusual fears, such as fear of elevators, stairs, toilets, balconies, vacuum cleaners, etc.	.531

*Note. FL = Factor Loads. ** = Inverted Items. *** = the Factors' (Subscales') Designation Used Below in the Text.*

Confirmatory factor analysis (CFA) conducted on the sample of children with ASD demonstrates good fit for the three-factor structure, particularly after allowing a residual correlation between items B2805 and N4601. Fit indices indicated acceptable model fit: χ^2 (295) = 656.203; CFI = .948; TLI = .943; RMSEA = .037 (p = 1.000); SRMR = .047. The positive residual correlation between these two items is explained by their shared reference to the child's food preferences, in addition to both belonging to Factor 3. Discriminant validity was supported by the Fornell-Larcker criterion (Fornell & Larcker, 1981). The minimum average variance extracted (AVE) was .217 (for Factor 3), while the highest inter-factor correlation was r = .31. The squared correlation (r^2 = .096) was lower than the minimum AVE, confirming discriminant validity. Internal consistency, assessed by Cronbach's α, was acceptable. Reliability coefficients ranged from .713 (for the Sensory Disintegration scale) to .783 (for the total 26-item scale).

Checking 3-Factor Structure Equivalence for Different Ages of the ASD Sample

Multigroup confirmatory factor analysis (MGCFA) was conducted to assess the equivalence of the three-factor structure across different subsamples. The results of the comparison between autistic boys and girls aged 1–7 years are presented in *Table 3*.

Table 3

Model Fit Indexes for 1–7-Year-Old Boys and Girls

Invariance testing	χ2	df	CFI	TLI	RMSEA
Configural	881.035	592	.958	.954	.033
Metric	923.686	615	.955	.953	.033
Scalar	946.704	638	.955	.955	.033
Strict	1011.689	664	.950	.951	.034

The unconstrained model demonstrates configurational equivalence of the factor structure across boys and girls, with excellent model fit: $\chi 2/df < 2$; CFI > .95; TLI > .95; RMSEA < .05. Differences in CFL, TLI, and RMSEA between the unconstrained model and subsequent constrained models did not exceed .01, indicating that measurement invariance held across all levels of constraint. Specifically, invariance was supported for: (a) metric invariance (equal factor loadings), (b) scalar invariance (equal intercepts), and (c) strict invariance (equal residual variances). These results confirm that the three-factor structure is equivalent across autistic boys and girls aged 1–7 years.

The results of the structures comparison for 1–2- and 3–7-year-old children are shown in *Table 4.*

Table 4

Model Fit Indexes for 1–2- and 3–7-Year-Olds

Invariance testing	χ2	df	CFI	TLI	RMSEA
Configural	119.434	590	.983	.982	.034
Metric	1465.855	613	.976	.975	.040
Scalar	1579.329	636	.974	.973	.041
Strict	1826.472	663	.968	.968	.045

The configurational equivalence of the structure for 1–2- and 3–7-year-old children is confirmed: $\chi 2/df < 2$; CFI > .95; TLI > .95; RMSEA < .05. The differences between CFI, TLI, and RMSEA for the previous and subsequent limited models do not reach .01. Thus, the equivalence of models for 1–2- and 3–7-year-olds is also certainly confirmed for all levels of restriction.

Age Differences on Scales for ASD Sample

Table 5 shows descriptive statistics for the compared groups, and *Figure 1* shows a graphical illustration.

According to the multidimensional Pillai Test, the effect of the Age factor is statistically significant ($p < .001$). According to one-dimensional ANOVA, the effects of the Age factor on Hyperactivity (Hyp) ($p = .027$; $\eta^2 = .008$) and Communication

Table 5

Descriptive Statistics of Scales for Groups Compared by Age and Diagnosis

Diagnosis		Age								
		1–2-year-olds			3–5-year-olds			6–7-year-olds		
		Hyp	Com	SD	Hyp	Com	SD	Hyp	Com	SD
ASD	Mean	.359	.589	.385	.294	.564	.427	.330	.477	.393
	N	130	130	130	487	487	487	300	300	300
	Std. Dev.	.280	.251	.255	.271	.256	.297	.273	.223	.224
DD	Mean	.418	.424	.350	.228	.269	.162	.231	.117	.136
	N	27	27	27	218	218	218	77	77	77
	Std. Dev.	.363	.292	.211	.255	.226	.165	.261	.150	.164
Total	Mean	.369	.561	.379	.273	.473	.345	.310	.404	.340
	N	157	157	157	705	705	705	377	377	377
	Std. Dev.	.295	.265	.248	.267	.282	.291	.273	.256	.237

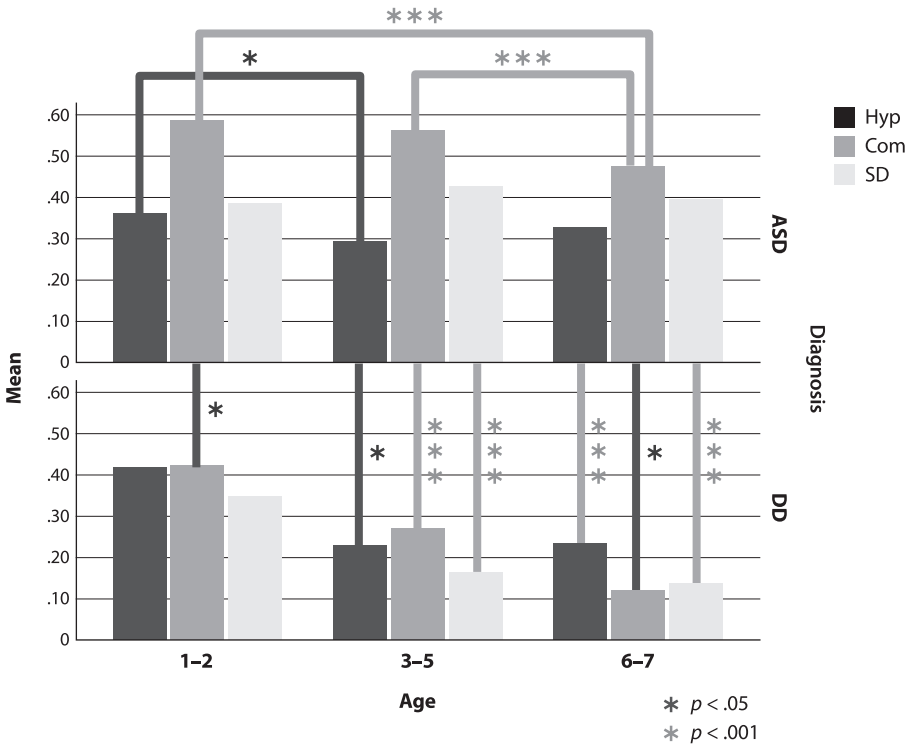

Figure 1. Age differences in the scales depending on the diagnosis (Y-axis is proportion of positive responses)

Disorders (Com) ($p < .001$; $\eta^2 = .031$) are statistically significant; the effects of the Age factor on Sensory Disintegration (SD) is not statistically significant ($p = .240$; $\eta^2 = .002$).

Hyperactivity (Hyp). According to the results of Post Hoc Comparisons, there are statistically significant differences only in the pair of 1–2- and 3–5-year-olds ($p = .04$; Cohen's $d = .238$): Hyperactivity in 1–2-year-old children is higher than in 3–5-year-olds. According to the contrast criterion, Hyperactivity in 1–2-year-old children is also higher than in 3–7-year-olds, but the differences do not reach statistical significance (t (914) = 1.806; $p = .071$).

Communication Disorders (Com). According to the results of Post Hoc Comparisons, there are statistically significant differences in the pair of 1–2- and 6–7-year-olds ($p < .001$; Cohen's $d = .455$), and in the pair of 3–5- and 6–7-year-olds ($p < .001$; Cohen's $d = .355$): the highest Communication Disorders are in 1–2-year-olds, the lowest are in 6–7-year-olds.

Differences in Scales Between Samples (ASD and DD) Depending on Age

Descriptive statistics are shown in *Table 6*. According to the multidimensional Pillai Test, the main effects of Age and Diagnosis factors, as well as the effect of their interaction, are statistically significant ($p < .001$). The primary focus of this study is the interaction between age and diagnosis (ASD vs. DD). While main effects would indicate overall group differences or general age-related changes, a significant interaction is required to determine whether the developmental trajectories of autistic symptoms differ between children with ASD and those with DD. Specifically, an interaction would suggest that the rate, direction, or pattern of change in symptoms across early childhood is not uniform across diagnostic groups — a finding with important implications for understanding the divergence of developmental pathways in the first years of life.

Hyperactivity (Hyp). The main effect of the Diagnosis factor is not statistically significant (F (1, 1233) = 2.232; $p = .135$; $\eta^2 = .014$). The effect of the interaction of Age and Diagnosis factors only slightly falls short of statistical significance (F(2, 1233) = 2.844; $p = .059$; $\eta^2 = .005$). *Table 6* shows the results of paired comparisons of averages.

Table 6

Post Hoc Comparisons – Age ∗ Diagnosis (Hyperactivity, Hyp)

		df	t	Cohen's d	p_{tukey}
(1–2 ASD)	(1–2 DD)	1233	–1.039	–.22	.905
	(3–5 DD)	1233	4.345	.481	< .001
	(6–7 DD)	1233	3.286	.473	.013
(3–5 ASD)	(3–5 DD)	1233	2.966	.242	.036
(6–7 ASD)	(6–7 DD)	1233	2.872	.367	.048

Note. P-value adjusted for comparing a family of six estimates

Hyperactivity rates were similarly elevated in 1–2-year-old children with ASD and DD, with no statistically significant differences between the two diagnostic groups. Across the entire age range, overall group differences in hyperactivity did not reach statistical significance. However, among children aged 3–5 years and 6–7 years, hyperactivity levels were significantly lower in the ASD group than in the DD group. This finding reflects a more pronounced age-related decline in hyperactivity among children with DD compared to those with ASD.

Communication Disorders (COM). The main effect of the Diagnosis factor is statistically significant (F (1, 1233) = 5.602; $p < .001$; $\eta^2 = .118$). The effect of the interaction of Age and Diagnosis factors is statistically significant (F (2, 1233) = 5.602; $p = .004$; $\eta^2 = .008$). *Table 7* shows the results of paired comparisons of averages.

Table 7

Post Hoc Comparisons – Age ∗ Diagnosis (Communication Disorders, COM)

		df	t	Cohen's d	p_{tukey}
(ASD 1–2)	(DD 1–2)	1233	3.276	.693	.014
(ASD 3–5)	(DD 3–5)	1233	15.204	1.239	< .001
(ASD 6–7)	(DD 6–7)	1233	11.85	1.514	< .001
(DD 1–2)	(DD 3–5)	1233	3.18	.649	.019
	(DD 6–7)	1233	5.764	1.289	< .001
(DD 3–5)	(DD 6–7)	1233	4.83	.64	< .001

Note. P-value adjusted for comparing a family of six estimates.

Across all three age ranges, Communication Disorders were significantly more pronounced in children with ASD than in those with DD. While Communication Disorders decreased with age in both diagnostic groups, the decline was significantly steeper among children with DD.

Table 8

Post Hoc Comparisons – Age ∗ Diagnosis (Sensory Disintegration, SD)

		df	t	Cohen's d	p_{tukey}
(ASD 1–2)	(DD 1–2)	1233	.666	.141	.986
(DD 1–2)	(DD 3–5)	1233	3.73	.761	.003
	(DD 6–7)	1233	3.871	.866	.002
(ASD 3–5)	(DD 3–5)	1233	13.168	1.073	< .001
(DD 3–5)	(DD 6–7)	1233	.791	.105	.969
(ASD 6–7)	(DD 6–7)	1233	8.143	1.04	< .001

Note. P-value adjusted for comparing a family of six estimates.

Sensory Disintegration (SD). The main effects of the Diagnosis factor (F (1, 1233) = 75.062; $p < .001$ $\eta^2 = .056$) and the Age factor (F (2, 1233) = 5.680; $p = .004$; $\eta^2 = .008$) are statistically significant. The effect of the interaction of Age and Diagnosis factors is statistically significant (F (2, 1233) = 8.634; $p < .001$; $\eta^2 = .013$). *Table 8* shows the results of paired comparisons of averages.

Sensory Disintegration was equally pronounced in 1–2-year-olds with ASD and DD. In contrast, among children aged 3–5 and 6–7 years, Sensory Disintegration declined significantly in the DD group, while no significant decline was observed in the ASD group. The main effect of diagnosis was twice as large for Communication Disorders as for Sensory Disintegration and approximately ten times as large for Communication Disorders as for Hyperactivity. Accordingly, Communication Disorders showed the largest differences between children with ASD and DD.

Discussion

The objectives of this study were twofold: first, to identify age-specific patterns of autism symptoms manifestation in younger, middle, and older preschoolers with ASD; and second, to compare these age-related trajectories with those observed in preschoolers with DD. Additionally, we tested the hypothesis that the three-factor structure would be equivalent across the samples of 1–2-year-olds and 3–7-year-olds with ASD.

The discussion proceeds as follows. First, we explain and provide theoretical substantiation for the three-factor structure, in which each factor formed a scale with adequate internal consistency (see *Table 2*). Second, we discuss age-related differences on these scales within the full ASD sample (*Table 5* and *Figure 1*). Finally, we examine the trajectories of symptom manifestation by comparing the ASD and DD samples across three age groups (*Tables 6–8*).

Interpretation of 3-Factor Structure of Autistic Symptoms in 1–2-Year-Olds

Exploratory factor analysis yielded a parsimonious three-factor structure comprising 26 items (8–9 per factor). The factors were interpreted as Communication Disorders (Com), Hyperactivity (Hyp) and Sensory Disintegration (SD). The first factor (Com) was named based on the items pertaining to the child's interaction with others. These items include: lack of interaction with others; non-use of communicative gestures; non-response to spoken speech; preference for solitary play; impaired joint attention within the context of interaction; and lack of imitation. The second factor was designated Hyperactivity (Hyp), as it comprises items related to excessive movements, noisiness, a tendency toward arousal and hyper-excitation with aggressive tendencies, difficulties in behavioral and motor control, and impulsivity. The third factor was labeled Sensory Disintegration (SD), as the items concern ritualistic behavior, altered sensory sensitivity, and insistence on sameness. Specific tasks include: preference of ritualism; excessive attachment to certain objects; and seeking or avoidance of particular sensory stimuli.

Previous research has demonstrated that markers and symptoms of ASD begin to emerge within the first year of life and can be detected as early as 6 to 18 months

of age (Tanner & Dounavi, 2021). In light of the three-factor structure obtained in the present study, it is noteworthy that both communication difficulties and sensory disintegration (or sensory processing difficulties) are core symptoms of autism. According to the DSM-5 (American Psychiatric Association, 2013), these symptoms typically emerge in early childhood and often persist across the lifespan (American Psychiatric Association, 2013). A recent meta-analysis revealed that communication differences in children on the autism spectrum may already be apparent by the age of one year (Collins et al., 2025). Researchers have identified several key indicators of communication difficulties at this age, including: absent or limited eye contact; failure to respond to one's name; absence of pointing gestures; difficulties with imitation (Frye et al., 2019); and very limited interest in playing with peers (McGlade et al., 2023). Additional characteristics include delayed speech onset, lack of gestural communication; difficulty understanding spoken language and atypical prosody (Lund et al., 2025). The classic autistic symptom of sensory disintegration can also manifest as early as 1–2 years of age. Affected children may exhibit repetitive movements, motor stereotypes, intense preoccupation with specific objects; unusual play with objects; and adherence to routine or insistence on sameness (Canu et al., 2021). Sensory disintegration is frequently associated with rituals and repetitive behaviors in autism (Boyd et al., 2010). Although repetitive behaviors in early childhood typically serve as a foundation for learning and exploration, in autism these behaviors may persist or intensify in response to ongoing sensory challenges. Rituals tied to sensory needs may become reinforcing because they provide consistent sensory input that the child craves or requires to feel comfortable (Donnellan, Hill & Leary, 2012). Given this established theoretical and empirical framework, the emergence of these two factors — Communication Disorders and Sensory Disintegration — within the factor structure of autism symptoms in 1–2-year-olds was not unexpected.

Unlike Communication Disorders and Sensory Disintegration, hyperactivity is not a core symptom of autism according to the DSM-5. Nevertheless, it represents a common and important feature to understand, as its presentation in autistic children often differs from that observed in typically developing children, children with developmental delay (DD), or those with attention-deficit/hyperactivity disorder (ADHD) alone. Research indicates that autistic children may exhibit constant movement such as running, climbing, jumping, or fidgeting, that can appear relentless and difficult to control (Al Ghamdi & AlMusailhi, 2024). They often shift focus rapidly between objects or activities, which impedes sustained engagement in play or learning. Additionally, sleep difficulties, including problems falling asleep or frequent night awakenings, may exacerbate daytime hyperactivity (Berenguer et al., 2024). Several potential mechanisms have been proposed to explain hyperactivity in this age group. First, hyperactivity may serve as a means of seeking sensory input or coping with sensory overload (Huang et al., 2024). Second, differences in brain maturation trajectories may affect activity levels and behavior regulation (O'Hearn & Lynn, 2023). Third, limited communication skills, particularly restricted verbal abilities, may lead children to express their needs or frustrations through increased activity or behavioral outbursts (Theodoratou, 2024). Finally, both over-stimulating and under-stimulating environments have been shown to exacerbate hyperactive behaviors (Gentil-Gutiérrez et al., 2021).

Thus, two of the three factors identified in the symptom structure for 1–2-year-olds with autism correspond to classic core symptoms of the condition. The third factor, Hyperactivity, represents a non-specific symptom that may nonetheless play a crucial role at this age, given its multifaceted nature and its susceptibility to influences such as sensory needs, neurological differences, and communication abilities.

When interpreting the three-factor structure of autism symptoms in preschoolers, it is essential to consider social and environmental factors. As autistic children grow, social interactions and environmental demands become increasingly complex, often rendering communication and sensory challenges more apparent. In the absence of early intervention and consistent support, these symptoms may remain stable or even intensify — a challenge that underscores the consequences of delayed autism diagnosis and its associated costs. It is noteworthy that our previous models of autistic symptoms in preschoolers yielded different factor structures when analyzed using structural equation modeling (SEM) on large samples of children, with the number of factors ranging from seven to eight, depending on age (Nasledov et al., 2023; Nasledov et al., 2024b). However, when we specifically examined age-related changes in autistic symptoms, we identified three clusters of correlated symptoms that closely resemble those described in the present manuscript (Nasledov et al., 2024a). Such variation in symptom combinations may be explained by differences in approaches and therapeutic strategies. Some children may develop compensatory strategies over time, even as core difficulties often persist. This pattern of stability in core symptom domains, coupled with variability in their surface-level expression, provides a basis for substantiating the equivalence of the three-factor structure across samples of 1–2-year-olds and 3–7-year-old boys and girls with ASD.

Age- and Diagnosis-Related Differences in Scales between ASD and DD Samples

Differences on the symptom scales between the ASD, DD, and normative sample were examined as a function of age and diagnosis (see *Tables 5–8* and *Figure 1*). The factor Age had a significant effect on Hyperactivity (Hyp) ($p = .027$; $\eta^2 = .008$) and Communication Disorders ($p < .001$; $\eta^2 = .031$). Hyperactivity was higher in 1–2-year-old children than in 3–5-year-olds, regardless of diagnosis. However, among 1–2-year-olds, hyperactivity was higher in children with DD than in their peers with ASD. Overall, hyperactivity in 1–2-year-old children was also higher than in 3–7-year-olds irrespective of diagnosis, though this difference did not reach statistical significance. As illustrated in *Figure 1*, the pattern of hyperactivity prevalence differs between diagnostic groups: children with DD show a distinct trajectory of declining hyperactivity with age, whereas children with ASD do not. Several factors may explain why hyperactivity is a common symptom shared by 1–2-year-olds with both ASD and DD. Overlapping neurodevelopmental and behavioral factors likely contribute (Martelli et al., 2025). Hyperactivity may serve as a means of seeking additional sensory input or as a reaction to sensory discomfort or overload. Although sensory issues are more characteristic of autism, children with DD may also experience difficulties processing sensory stimuli, which can contribute to hyperactive behaviors (Delgado-Lobete

et al., 2020). Additionally, lower cognitive control in 1–2-year-olds may affect their ability to understand and follow instructions, resulting in hyperactive or impulsive actions (Peisch & Arnett, 2024). More broadly, at this age, the brain systems underlying attention and behavioral regulation are still maturing across all neurodevelopmental pathways; thus, hyperactivity may represent an exaggerated expression of a normative developmental feature in both ASD and DD. However, our results show that children with DD exhibit a distinct trajectory of declining hyperactivity across the preschool years, in contrast to the ASD sample. It is possible that children with temporary delays, whether due to environmental factors or responsive to targeted interventions, may "catch up," with hyperactivity diminish as they develop (Wootton et al., 2022). In contrast, among preschoolers with ASD, hyperactivity remains highly prevalent, with both overlaps and distinctions in executive function impairment (Christoforou et al., 2023; Esposito et al., 2024).

Communication Disorders emerged as the most prominent symptom group in the ASD sample across all ages. Although a gradual decline in communication difficulties was observed over time in autistic children, this symptom domain remained the most distinguishing feature between the ASD and DD samples. In the DD sample, communication problems in 1–2-year-olds were nearly as pronounced as in children with ASD of the same age, but tended to resolve over time. This finding aligns with previous research. It is well known that autistic children struggle with developing communication skills throughout the preschool years (Kauley et al., 2024), whereas children with DD may overcome problems in communication due to brain maturation and the acquisition of behavioral control by older preschool age (Sheridan et al., 2019).

Sensory Disintegration was significantly more pronounced in children with ASD aged 3–5 years and 6–7 years compared to their peers with DD. However, at 1–2 years of age, the difference between the two diagnostic groups was less pronounced. A clear developmental dynamic was observed: sensory issues declined over time in children with DD, whereas they remained prevalent across the entire preschool period in children with ASD. These findings align with previous research. Sensory issues are considered a core symptom of autism and are consistently reported in children with ASD (Khawer et al., 2025). While young children with DD are also known to experience distinct sensory challenges across different sensory systems (Rodríguez-Armendariz, Vela-Romero & Galiana, 2024), this group appears to benefit from timely targeted intervention, with positive outcomes in overcoming sensory difficulties (Liu et al., 2025).

Overall, when comparing both samples, the main effect of the diagnosis was most pronounced for Communication Disorders. This symptom domain appears to be the key differentiator between children with ASD and those with DD. However, it should be noted that while Communication Disorders represent a highly salient and often early-presenting feature that helps distinguish ASD from DD in preschoolers, it is the specific nature of these communication difficulties, particularly deficits in social communication, combined with restricted and repetitive behaviors and interests that ultimately leads to a diagnosis of ASD. A holistic approach that considers the comprehensive clinical picture, rather than any single symptom in isolation, is essential to guiding the diagnostic process.

Conclusion

The findings of this study demonstrate that symptoms of Communication Disorders, Sensory Disintegration, and Hyperactivity persist throughout the preschool years in children with autism spectrum disorder (ASD). We interpret this persistence as reflecting fundamental and stable neurodevelopmental differences that affect brain function, sensory processing, and self-regulation. Although early intervention can improve functioning and reduce symptom severity, these challenges often continue across the preschool period due to the complexity and stability of the underlying neurological mechanisms. Two of the three factors identified in the symptom structure for 1–2-year-olds with autism correspond to classic core symptoms of the condition. The third factor, hyperactivity, represents a non-specific symptom that may nevertheless play a crucial role at this age, given its multifaceted nature and its susceptibility to influences such as sensory needs, neurological differences, and communication abilities. Hyperactivity emerged as a shared symptom for 1–2-year-olds with both ASD and DD. This overlap likely arises from shared neurodevelopmental delays affecting brain regions responsible for attention, motor control, and sensory processing. Difficulties in self-regulation, communication, and sensory integration contribute to increased motor activity and impulsivity in both groups. Although the underlying mechanisms may differ in detail, the behavioral manifestation of hyperactivity is common due to overlapping developmental challenges at this age. However, a key divergence was observed: children with DD tend to overcome hyperactivity over preschool age, whereas in children with ASD it remains pronounced. A similar pattern was observed for Sensory Disintegration. This symptom was also shared between diagnostic groups at age 1–2 years, with a distinct developmental trajectory toward resolution in children with DD, while remaining prevalent in children with ASD across the preschool period. Notably, Communication Disorders emerged as the most distinct symptom for differentiating children with ASD from those with DD, regardless of age. This finding underscores the centrality of social communication deficits in autism and highlights their diagnostic utility, particularly when considered alongside other core and associated symptoms.

Limitations

Several limitations of this study should be acknowledged. The sample size of children with ASD and DD was limited. This constraint stems from the fact that 1–2-year-olds do not yet attend preschool educational institutions, and are assessed much less frequently than older children, making recruitment for this age group inherently challenging. A second limitation concerns the number of symptom groups identified. While the present study yielded a three-factor structure, our previous research has demonstrated that the number of relatively independent symptom groups in preschoolers with ASD can reach seven (for 3–4-year-olds) and even eight (for 5–7-year-olds). The identification of only three symptom groups in this study can be attributed to two factors. First, at 1–2 years of age, ASD symptoms are not yet as differentiated

as they are in older preschoolers — a finding that is developmentally expected. Second, to identify age- and sex-related specifics of symptom manifestation in 1–2-year-olds, it was necessary to establish a structure that could be meaningfully compared with that of older preschoolers. Importantly, the three symptom groups identified — Communication Disorders, Hyperactivity, and Sensory Disintegration — are indeed common across the entire age range from 1 to 7 years.

Another limitation, consistent with our previous studies, is that the symptoms identified were those detected and noted by specialists working directly with each child. Although these specialists were familiar with the children's typical behavior due to regular interactions, the full range of objective manifestations of ASD may be considerably broader than captured in the present assessment.

Practical Implications

Understanding age-related changes in autistic symptoms during the preschool years carries important practical implications for clinical, educational, and caregiving settings. Recognizing that autistic symptoms may manifest differently depending on the child's age can help practitioners identify autism more accurately and at an earlier stage. Revealing developmental trajectories of symptom manifestation is crucial for differentiating children with ASD from those with DD, as understanding how these symptoms typically evolve over time is key to accurate diagnosis. Knowledge of typical developmental trajectory allows professionals to determine whether a child is simply lagging behind in certain areas or is exhibiting a deviant pattern of development characteristic of ASD. The earlier a child receives an ASD diagnosis, the sooner appropriate interventions can be initiated, and the better the prognosis for long-term development. Practitioners should therefore be particularly attentive to Communication Disorders in 1–2-year-olds as the most salient early sign that may indicate underlying neurological differences associated with autism. In contrast, Hyperactivity and Sensory Disintegration at this age are shared symptoms that may be present regardless of diagnosis and should be interpreted within the broader clinical context.

Ethics Statement

This study was approved by the St. Petersburg State University Ethics Committee (Approval No. 02-285, RK No. 115-03-27, dated October 10, 2024). Prior to participation, parents provided written informed consent for both their own participation in the survey and their children's involvement in the online assessment. The consent forms provided clear and comprehensive information regarding the study's objectives, procedures, and potential benefits. Considering that 1- and 2-year-old autistic children are particularly sensitive to environmental changes, the study was designed to minimize all physical and psychological risks. To ensure a comfortable environment and prevent distress, pain, or sensory overload, data were collected via online questionnaires completed by specialists who work with the children on a permanent basis. This approach ensured that no unfamiliar procedures were introduced, there-

by maintaining the children's routine and emotional well-being. The research group guarantees that data is collected in an anonymized form, with a unique code assigned to each participant. The group also assumes responsibility for the confidentiality of the collected data and its secure storage in protected databases accessible only to the research team. Only anonymized data will be used in publications and reports. To encourage participation, investigators offered direct benefits to the children through enhanced interventions tailored to the results of their data collection. Additionally, the study contributed to the broader autistic community by providing free, open access to practical research findings via the project's electronic portal. Parents were explicitly informed of their right to withdraw their child or themselves from the study at any time without negative consequences. Although the procedures posed no risk of distress, the voluntary nature of participation was strictly maintained to ensure ethical compliance. During parental surveys, researchers remained highly sensitive to family privacy, ensuring that questioning was non-intrusive. The data collection was conducted by experts with extensive experience in working with young autistic children. This specialized expertise was critical in avoiding common errors associated with standardized tools like the ADOS-2. In typical clinical settings, lack of prior rapport between an observer and a child can cause frustration, leading to atypical behavior and subsequent data bias. By utilizing researchers familiar with the participants, the study ensured a low-pressure environment that captured more accurate behavioral observations.

Informed Consent from the Participants' Legal Guardians (if the participants were minors)

Written informed consent to participate in this study was provided by the participants' legal guardian/next of kin.

Authors' Contributions

Conceptualization and methodology, A.N.; software, S.M.; validation and formal analysis, A.N.; investigation and resources, S.M.; data curation, A.N.; writing and original draft preparation, L.T.; writing review and editing, L.T.; visualization, A.N.; supervision, A.N.; project administration, A.N.; funding acquisition, A.N. All authors discussed the results and contributed to the final manuscript.

Conflict of Interest

The authors declare no conflict of interest.

Funding

The reported study was funded by the Russian Science Foundation, project No. 23-18-00155-P. "Study of predictive indicators of autism spectrum disorders in preschool children" (https://rscf.ru/en/project/23-18-00155-П/).

References

Al-Beltagi, M. (2021). Autism medical comorbidities. *World Journal of Clinical Pediatrics, 10*(3), 15–28. https://doi.org/10.5409/wjcp.v10.i3.15

Al Ghamdi, K., & AlMusailhi, J. (2024). Attention-deficit hyperactivity disorder and autism spectrum disorder: Towards better diagnosis and management. *Medical Archives, 78*(2), 159–163. https://doi.org/10.5455/medarh.2024.78.159-163

American Psychiatric Association (2013). *Diagnostic and statistical manual of mental disorders* (5th ed.). https://doi.org/10.1176/appi.books.9780890425596

Berenguer, C., Baixauli, I., Rosa, E., & De Stasio, S. (2024). Sleep problems in children with autism spectrum disorder and attention-deficit/hyperactivity disorder: A comparative study and effects on communication skills. *Autism Research, 17*(2), 355–365. https://doi.org/10.1002/aur.3077

Boyd, B.A., Baranek, G.T., Sideris, J., Poe, M.D., Watson, L.R., Patten, E., & Miller, H. (2010). Sensory features and repetitive behaviors in children with autism and developmental delays. *Autism Research, 3*(2), 78–87. https://doi.org/10.1002/aur.124

Byrne, B.M. (2010). *Structural equation modeling with AMOS: Basic concepts, applications and programming* (2nd ed.). Routledge.

Canu, D., Van der Paelt, S., Canal-Bedia, R., Posada, M., Vanvuchelen, M., & Roeyers, H. (2021). Early non-social behavioural indicators of autism spectrum disorder (ASD) in siblings at elevated likelihood for ASD: A systematic review. *European Child & Adolescent Psychiatry, 30*(4), 497–538. https://doi.org/10.1007/s00787-020-01487-7

Chen, F.F. (2007). Sensitivity of goodness of fit indexes to lack of measurement invariance. *Structural Equation Modeling, 14*, 464–504. https://doi.org/10.1080/10705510701301834

Christoforou, M., Jones, E.J.H., White, P., & Charman, T. (2023). Executive function profiles of preschool children with autism spectrum disorder and attention-deficit/hyperactivity disorder: A systematic review. *JCPP Advances*, *3*(1), e12123. https://doi.org/10.1002/jcv2.12123

Collins, I.M., Halter, F., Schächinger Tenés, L., Lieb, R., & Meyer, A.H. (2025). A meta-analysis of applied behavior analysis-based interventions to improve communication, adaptive, and cognitive skills in children on the autism spectrum. *Review Journal of Autism and Developmental Disorders*, 1–21. https://doi.org/10.1007/s40489-025-00506-0

Delgado-Lobete, L., Pértega-Díaz, S., Santos-Del-Riego, S., & Montes-Montes, R. (2020). Sensory processing patterns in developmental coordination disorder, attention deficit hyperactivity disorder and typical development. *Research in Developmental Disabilities*, *100*, 103608. https://doi.org/10.1016/j.ridd.2020.103608

Donnellan, A.M., Hill, D.A., & Leary, M.R. (2012). Rethinking autism: Implications of sensory and movement differences for understanding and support. *Frontiers in Integrative Neuroscience, 6*, 124. https://doi.org/10.3389/fnint.2012.00124

Esposito, D., Bernardi, K., Belli, A., Gasparri, V., Romano, S., Terenzi, L., Zanatta, M.E., Iannotti, S., & Ferrara, M. (2024). The hidden impact of ADHD symptoms in preschool children with autism: Is there a link to somatic and sleep disorders? *Behavioral Sciences, 14*(3), 211. https://doi.org/10.3390/bs14030211

Fornell, C., & Larcker, D.F. (1981). Evaluating structural equation models with unobservable variables and measurement error. *Journal of Marketing Research, 18*, 39–50. https://doi.org/10.2307/3151312

Frye, R.E., Vassall, S., Kaur, G., Lewis, C., Karim, M., & Rossignol, D. (2019). Emerging biomarkers in autism spectrum disorder: A systematic review. *Annals of Translational Medicine, 7*(23), 792. https://doi.org/10.21037/atm.2019.11.53

Fulceri, F., Caruso, A., Micai, M., Gila, L., Tancredi, R., Fatta, L.M., Galati, G., d'Amico, R., Schünemann, H.J., Balduzzi, S., Cinquini, M., & Scattoni, M.L. (2025). Autism diagnosis in children and adolescents: A systematic review and meta-analysis of test accuracy. *Neuroscience and Biobehavioral Reviews, 173,* 106164. https://doi.org/10.1016/j.neubiorev.2025.106164

Gentil-Gutiérrez, A., Cuesta-Gómez, J.L., Rodríguez-Fernández, P., & González-Bernal, J.J. (2021). Implication of the sensory environment in children with autism spectrum disorder: Perspectives from school. *International journal of environmental research and public health, 18*(14), 7670. https://doi.org/10.3390/ijerph18147670

Green, S.B. (1991). How many subjects does it take to do a regression analysis. *Multivariate Behavioral Research, 26*(3), 499–510. https://doi.org/10.1207/s15327906mbr2603_7

Gitimoghaddam, M., Chichkine, N., McArthur, L., Sangha, S.S., & Symington, V. (2022). Applied behavior analysis in children and youth with autism spectrum disorders: A scoping review. *Perspectives on Behavior Science, 45*(3), 521–557. https://doi.org/10.1007/s40614-022-00338-x

Huang, Z., Wang, F., Xue, L., Zhu, H., & Zou, X. (2024). Relationships between sensory processing and executive functions in children with combined ASD and ADHD compared to typically developing and single disorder groups. *Brain sciences, 14*(6), 566. https://doi.org/10.3390/brainsci14060566

Issac, A., Halemani, K., Shetty, A., Thimmappa, L., Vijay, V.R., Koni, K., Mishra, P., & Kapoor, V. (2025). The global prevalence of autism spectrum disorder in children: a systematic review and meta-analysis. *Osong public health and research perspectives, 16*(1), 3–27. https://doi.org/10.24171/j.phrp.2024.0286

Kauley, N., John, J.R., Barr, K.R., Wu, W.T., Grove, R., Masi, A., & Eapen, V. (2024). Predicting communication skills outcomes for preschool children with autism spectrum disorder following early intervention. *Neuropsychiatric Disease and Treatment*, *20*, 35–48. https://doi.org/10.2147/NDT.S435740

Khawer, S., Ismail, M. & Zehra, S. (2025). Linking sensory processing challenges to autism-related factors: Observational evidence from Pakistan. *Middle East Current Psychiatry, 32*, 37. https://doi.org/10.1186/s43045-025-00533-x

Kline, R.B. (2011). *Principles and practice of structural equation modeling* (3rd ed.). The Guilford Press.

Li, C.-H. (2016). The performance of ML, DWLS, and ULS estimation with robust corrections in structural equation models with ordinal variables. *Psychological Methods, 21*(3), 369–387. https://doi.org/10.1037/met0000093

Liu, S., Zhu, M., Yi, C., & Zeng, D. (2025). Early rehabilitation interventions for global developmental delay in children: A narrative review. *Frontiers in Pediatrics*, *13*, 1576324. https://doi.org/10.3389/fped.2025.1576324

Lund, S.K., Quach, W., Weissling, K., & McKelvey, M. (2025). Augmentative and alternative communication assessment for children on the autism spectrum: Protocol development and content validation. *Language, speech, and hearing services in schools, 56*(1), 1–16. https://doi.org/10.1044/2024_LSHSS-23-00127

Nasledov, A., Miroshnikov, S., Tkacheva, L., Miroshnik, K., & Semeta, M.U. (2021). Application of psychometric approach for ASD evaluation in Russian 3–4-year-olds. *Mathematics, 9*(14), 1667. https://doi.org/10.3390/math9141608

Nasledov, A., Miroshnikov, S., Tkacheva, L., & Fedorov, S. (2023). Factor structure of ASD symptoms in Russian 3–4-year-olds. *OBM Neurobiology, 7*(4), 1–28. https://doi.org/10.21926/obm.neurobiol.2304190

Nasledov, A., Tkacheva, L., & Miroshnikov, S. (2024a). Factor structure of autism symptoms in 5–6-year-old children: Age perspective. *OBM Neurobiology, 8*(4), 259. https://doi.org/10.21926/obm.neurobiol.2404259

Nasledov, A., Tkacheva, L., & Miroshnikov, S. (2024b). 7-factor model of autism symptoms in 3–6-year-olds: Age-related changes. *Terapia Psicológica* [Psychological Therapy], *42*(3), 379–414. https://doi.org/10.4067/s0718-48082024000300379

Maksimović, S., Marisavljević, M., Stanojević, N., Ćirović, M., Punišić, S., Adamović, T., Đorđević, J., Krgović, I., & Subotić, M. (2023). Importance of early intervention in reducing autistic symptoms and speech-language deficits in children with autism spectrum disorder. *Children, 10*(1), 122. https://doi.org/10.3390/children10010122

Masi, A., DeMayo, M.M., Glozier, N., & Guastella, A.J. (2017). An overview of autism spectrum disorder, heterogeneity and treatment options. *Neuroscience Bulletin, 33*(2), 183–193. https://doi.org/10.1007/s12264-017-0100-y

Martelli, M.E., Gigliotti, F., Giovannone, F., Lentini, G., Manti, F., & Sogos, C. (2025). Developmental patterns in autism and other neurodevelopmental disorders in preschool children. *Children*, *12*(2), 125. https://doi.org/10.3390/children12020125

McGlade, A., Whittingham, K., Barfoot, J., Taylor, L., & Boyd, R.N. (2023). Efficacy of very early interventions on neurodevelopmental outcomes for infants and toddlers at increased likelihood of or diagnosed with autism: A systematic review and meta-analysis. *Autism Research, 16*(6), 1145–1160. https://doi.org/10.1002/aur.2924

Mualem, R., Morales-Quezada, L., Farraj, R.H., Shance, S., Bernshtein, D.H., Cohen, S., Mualem, L., Salem, N., Yehuda, R.R., Zbedat, Y., Waksman, I., & Biswas, S. (2024). Econeurobiology and brain development in children: Key factors affecting development, behavioral outcomes, and school interventions. *Frontiers in Public Health*, *12*, 1376075. https://doi.org/10.3389/fpubh.2024.1376075

O'Hearn, K., & Lynn, A. (2023). Age differences and brain maturation provide insight into heterogeneous results in autism spectrum disorder. *Frontiers in human neuroscience, 16*, 957375. https://doi.org/10.3389/fnhum.2022.957375

Peisch, V., & Arnett, A.B. (2024). Neural activation, cognitive control, and attention deficit hyperactivity disorder: Evaluating three competing etiological models. *Development and Psychopathology*, *36*(1), 255–265. https://doi.org/10.1017/S095457942200116X

Rodríguez-Armendariz, E., Vela-Romero, M., & Galiana, A. (2024). Sensory processing challenges in children with neurodevelopmental disorders and genetic conditions: An observational study. *NeuroSci, 5*(3), 339–353. https://doi.org/10.3390/neurosci5030027

Saldaris, J., Leonard, H., Wong, K., Jacoby, P., Spence, M., Marsh, E.D., Benke, T.A., Demarest, S., & Downs, J. (2024). Validating the communication and symbolic behavior scales-developmental profile infant-toddler checklist (CSBS-DP ITC) beyond infancy in the CDKL5 deficiency disorder. *Journal of Autism and Developmental Disorders, 54*(7), 2526–2535. https://doi.org/10.1007/s10803-023-06002-w

Sheridan, S.M., Knoche, L.L., Boise, C.E., Moen, A.L., Lester, H., Edwards, C.P., Schumacher, R., & Cheng, K. (2019). Supporting preschool children with developmental concerns: Effects of the Getting Ready intervention on school-based social competencies and relationships. *Early Childhood Research Quarterly, 48,* 303–316. https://doi.org/10.1016/j.ecresq.2019.03.008

Tanner, A., & Dounavi, K. (2021). The emergence of autism symptoms prior to 18 months of age: A systematic literature review. *Journal of Autism and Developmental Disorders, 51*, 973–993. https://doi.org/10.1007/s10803-020-04618-w

Theodoratou, M. (2024). Communication issues in co-occurring ADHD and autism spectrum disorders. Evaluative approaches and targeted interventions: Mini review. *Postepy Psychiatrii Neurologii, 33*(3), 188–195. https://doi.org/10.5114/ppn.2024.145214

Thurman, A.J., Nunnally, A.D., Nguyen, V., Berry-Kravis, E., Sterling, A., Edgin, J., Hamilton, D., Aschkenasy, J., & Abbeduto, L. (2024). Short-term and long-term stability of the Autism Diagnostic Observation Schedule (ADOS-2) calibrated comparison scores (CCS) and classification scores in youth with down syndrome or fragile X syndrome with intellectual disability. *Journal of Autism and Developmental Disorders*. Advance online publication. https://doi.org/10.1007/s10803-024-06535-8

Waizbard-Bartov, E., & Miller, M. (2023). Does the severity of autism symptoms change over time? A review of the evidence, impacts, and gaps in current knowledge. *Clinical Psychology Review, 99*, 102230. https://doi.org/10.1016/j.cpr.2022.102230

Wootton, R.E., Riglin, L., Blakey, R., Agnew-Blais, J., Caye, A., Cadman, T., Havdahl, A., Gonçalves, H., Menezes, A.M.B., Wehrmeister, F.C., Rimfeld, K., Davey Smith, G., Eley, T.C., Rohde, L.A., Arseneault, L., Moffitt, T.E., Stergiakouli, E., Thapar, A., & Tilling, K. (2022). Decline in attention-deficit hyperactivity disorder traits over the life course in the general population: Trajectories across five

population birth cohorts spanning ages 3 to 45 years. *International Journal of Epidemiology*, *51*(3), 919–930. https://doi.org/10.1093/ije/dyac049

Zheng, R.M., Chan, S.P., Law, E.C., Chong, S.C., & Aishworiya, R. (2024). Validity and feasibility of using the Modified Checklist for Autism in Toddlers, R evised with Follow-Up (M-CHAT-R/F) in primary care clinics in Singapore. *Autism, 28*(7), 1758–1771. https://doi.org/10.1177/13623613231205748

Original manuscript received June 2, 2026
Revised manuscript accepted August 27, 2026
First published online September 1, 2026

To cite this article: Nasledov, A.D., Tkacheva, L.O., Miroshnikov, S.A. (2026). Age-Related Trajectories in Symptom Manifestation in Preschoolers with Autism and Developmental Delay. *Psychology in Russia: State of the Art*, *19*(3), 3–24. DOI: 10.11621/pir.2026.0301

Psychology in Russia: State of the Art
Volume 19, Issue 3, 2026

Russian Psychological Society
Lomonosov Moscow State University

ORGANIZATIONAL PSYCHOLOGY

Working from Home Across National Contexts: Demands, Resources, Engagement, and Productivity Among Italian and Russian Employees

Ferdinando Toscano[a], Salvatore Zappalà[b*],
Marina V. Polevaya[c], Elena V. Kamneva[c]

[a] *Universitas Mercatorum, Rome, Italy*
[b] *Alma Mater Studiorum University of Bologna, Italy*
[c] *Financial University under the Government of the Russian Federation, Moscow, Russia*

*Corresponding author. E-mail: salvatore.zappala@unibo.it

Background. The expansion of work from home has changed the material, social, and self-regulatory conditions under which employees organize their work activities. Although remote work can support flexibility and concentration, its implications for productivity depend on the balance between work-related demands and resources.

Objective. Drawing on the Job Demands–Resources (JD-R) model, this study examined whether environmental distractions, job autonomy, and self-leadership are related to perceived job productivity among Italian and Russian remote workers, and whether remote work stress and remote work engagement mediate these relationships.

Design. Data were collected through an online questionnaire administered to employees working from home in Italy and Russia. A total of 493 respondents provided complete data for all the variables included in the research model and were retained for the analyses. Confirmatory factor analyses, measurement invariance, structural equation modeling, and multigroup analyses were conducted.

Results. Environmental distractions were associated with higher remote work stress and lower perceived job productivity. Job autonomy and self-leadership were positively related to remote work engagement, which in turn predicted higher productivity. Unexpectedly, both resources showed negative direct associations with productivity, although the effect of self-leadership was weak. Multigroup analyses indicated that the overall pattern of relationships was broadly consistent across the Italian and Russian samples.

Keywords:
remote work, work from home, job demands-resources model, environmental distractions, self-leadership, job productivity, cross-national psychology

© Toscano, F., Zappalà, S., Polevaya, M.V., Kamneva, E.V., 2026

ISSN 2074-6857 (Print) / ISSN 2307-2202 (Online)
http://psychologyinrussia.com

Conclusion. The findings support the motivational pathway of the JD-R model in remote work settings, while also suggesting that autonomy and self-leadership may become ambivalent under high self-organization demands. They also indicate that key JD-R mechanisms operate similarly across the Italian and Russian samples, although some cross-national differences remain.

Introduction

Remote work has become a major form of work organization, reshaping where, when, and how employees perform their tasks. This transformation is not merely a change in physical location. It also modifies the material environment in which work is carried out, the opportunities for direct supervision and social interaction, and the degree to which employees regulate their own job tasks. While remote work can provide flexibility and reduce some office-based interruptions, it can also introduce new demands, such as inadequate workspaces, household noise, blurred boundaries, and reduced structure in daily work routines (Galanti et al., 2021; Xiao et al., 2021). Accordingly, the effect of working from home on job productivity remains debated. Some studies suggest that remote work can improve performance, whereas others show that productivity may decline under specific individual, organizational, or environmental conditions (Allen et al., 2015; Monteiro et al., 2021; Toscano & Zappalà, 2020a; Troll et al., 2022). These inconsistencies indicate the need for more fine-grained analyses of the demands and resources that shape productivity in home-based work. Although the data were collected during the COVID-19 pandemic, when employees were undergoing an emergency transition to home-based work, the study remains relevant for contemporary remote and hybrid work because it examines psychological mechanisms that continue to shape distributed work arrangements. Recent evidence suggests that remote and hybrid work are neither uniformly beneficial nor detrimental: their outcomes depend on how flexibility, work design, health-related demands, and employees' self-regulatory capacities are combined (Bloom et al., 2024; Dong et al., 2025; Wells et al., 2023). Therefore, examining how demands and resources operate in home-based work can contribute to current debates on sustainable and effective flexible work design.

This study addresses this issue by adopting the Job Demands–Resources (JD-R) model (Bakker & Demerouti, 2007, 2017) to examine whether specific characteristics of remote work are related to employees' perceived job productivity. In the JD-R model, job demands are physical, psychological, social or organizational aspects of the job that require sustained effort and are associated with psychological costs; job and personal resources, instead, refer to physical, psychological, social or organizational aspects of the job that help the employee to achieve work goals and stimulate personal growth and learning and development. In the context of work from home, we conceptualize environmental distractions as a "job demand" that may increase

stress and undermine productivity. Noise, inadequate furniture, limited work area size, wall colors, and cleanliness of the work environment can become distracting factors (Lee & Brand, 2005) that generate strain (Wineman & Barnes, 2018) and may initiate a health-impairment process. By contrast, job autonomy is conceptualized as a "job resource", because it can satisfy a primary motivational need for control and self-determination (Deci et al., 2017). Self-leadership, defined as a set of behavioral and cognitive strategies aimed at self-regulation and goal attainment (Stewart et al., 2011), is considered a "personal resource" that may further support productivity in remote work.

The cross-national design of this study is central to its contribution. Data were collected from Italian and Russian remote workers, to examine whether JD-R mechanisms operate similarly across two socio-organizational contexts that differ in work traditions, managerial assumptions, and previous familiarity with autonomous remote work practices. Rather than treating national context merely as a background condition, we use the Italian and Russian samples to provide an exploratory check of whether the proposed psychological mechanisms show a similar pattern across different contexts. This orientation is particularly relevant for research on work psychology and ergonomics. From the activity theory perspective, rooted in the Russian psychological tradition (Maidansky, 2021), job performance depends on how work activity is organized, including task demands, available tools, environmental constraints, methods of action, and self-regulatory processes (Bedny & Karwowski, 2006; Bedny & Meister, 1997). Accordingly, when working from home, employees must organize tasks, regulate attention, manage environmental constraints, and sustain motivation in a setting that is only partially designed for formal work. Thus, the study has two aims: first, to integrate environmental distractions, job autonomy, and self-leadership into a unified JD-R framework on remote work productivity; second, to examine whether the pattern of associations shows convergence among Italian and Russian employees.

Environmental Distractions, Remote Work Stress, and Job Productivity

In this study, we investigate the relationship between environmental distractions experienced when working from home and employees' perceived job productivity.

Environmental distractions are conceptualized as undesirable stimuli that cause employees irritation and negative experiences (Lee & Brand, 2005). They create attentional conflict by diverting attention from work tasks, thereby impairing cognitive control and increasing the psychological cost of work. Consistent with the JD-R framework, in this study distractions are considered as job demands that consume cognitive and emotional resources and contribute to a health-impairment process.

Job productivity, while complex to define, can be understood as the ratio between the value of output produced and the psychological and physical costs of input (Murphy, 1990). In organizational settings, productivity refers to the successful

completion of work tasks and responsibilities. In this study, we adopt a subjective measure of job productivity, capturing employees' perceptions of the quality of their output when working remotely compared to their usual productivity when working in-person.

Previous research has documented that environmental distractions (such as family interruptions, inadequate workspace, or household noise) reduce perceived control, increase effort, and heighten employee feelings of frustration and inefficacy (Keller et al., 2020; Xiao et al., 2021). Moreover, frequent interruptions may require employees to switch tasks or change physical locations (*e.g.*, working at a dinner table that must be cleared), increasing time pressure and cognitive load (Baethge & Rigotti, 2013) that are typical causes of job stress.

Job stress is a condition that develops when the interaction between individuals and their work environment deviates from its normal functioning, resulting in the accumulation of negative emotions and cognitive strains generated by work demands (Weinert et al., 2015). Accordingly, we propose that environmental distractions are negatively related to job productivity, and that remote work stress, defined as the stress experienced because of the challenges of working remotely (Toscano & Zappalà, 2020b), acts as a mediator. In fact, the JD-R model suggests that job demands, when ineffectively managed, may initiate a health-impairment process that ultimately impacts job performance. Thus, we hypothesize:

H1a: Environmental distractions are negatively related to job productivity.

H1b: Remote work stress mediates the relationship between environmental distractions and job productivity.

Job Autonomy, Self-Leadership, Remote Work Engagement, and Job Productivity

In parallel with the consideration of the health-impairment process, this study also investigates the effect of job and personal resources on job productivity, considering the motivational process postulated in the JD-R theory (Bakker & Demerouti, 2007, 2017). We focus on two resources, job autonomy and self-leadership, and their capacity to enhance remote work engagement defined as a positive, fulfilling, work-related state of mind experienced during remote work (Memon et al., 2021; Toscano & Zappalà, 2021).

Job autonomy refers to the extent to which employees can make decisions on how to complete their job tasks or how to modify or influence their own workplaces (Morgeson & Humphrey, 2006). Job autonomy has been long thought to be an advantage of telework, since it allows employees to choose where, when, and how to work, providing them with independence, discretion, and freedom in scheduling work and determining the procedures to be used (Allen et al., 2015). Furthermore, job autonomy leads to an increase of confidence in the job performance in a specific task (Saragih, 2011), another element supporting the existence of a positive relationship between job autonomy and job productivity. Hence, we posit that:

H2a: Job autonomy is positively related to job productivity.

The JD-R model suggests that job resources promote work engagement, which in turn leads to higher performance (Bakker & Demerouti, 2007, 2017). Tasks that allow greater autonomy tend to foster engagement by satisfying employees' psychological needs for competence and control (Sonnentag, 2017). Work engagement itself is characterized by vigor, dedication, and absorption (Bakker & Demerouti, 2024; Memon et al., 2021) and has been consistently associated with higher job performance (Bakker & Demerouti, 2024). Therefore, we posit:

H2b: Remote work engagement mediates the relationship between job autonomy and job productivity.

Alongside job autonomy, self-leadership is here investigated as a personal resource. Self-leadership is defined as the process through which individuals influence and motivate themselves to achieve work goals using behavioral and cognitive strategies (Stewart et al., 2011), which in the present study are goal setting and self-monitoring.

Goal setting is the practice of defining specific and challenging goals that provide direction and motivation for task accomplishment (Locke & Latham, 2013). From a motivational perspective, goal setting acts as an action plan that influences the amount of effort, direction, and persistence of a certain action. Goals represent what people aim to accomplish, so they help employees to direct their energies in order to move, step by step, towards task completion (Troll et al., 2022).

Self-monitoring is defined as the active observation of one's own work, and such strategy leads to an awareness of when and why employees engage in behaviors that are functional to their goals (Stewart et al., 2011). Previous research found a positive correlation between self-monitoring behaviors and job performance (Moser & Galais, 2007). Furthermore, previous studies found that self-leadership provides an increased focus on cognitive processes, leading to enhanced job performance (Stewart et al., 2011), and prevents negative functional states, sustaining work efficiency under demanding conditions (Leonova, 2010). For these reasons, we posit that:

H3a: Self-leadership is positively associated with job productivity.

We consider self-leadership as a personal resource and recognize the existence of a growing body of evidence highlighting the positive impact of self-leadership on work engagement (Breevaart et al., 2016). Working with clear goals and monitoring each step toward them streamlines work processes and informs workers about their progress, fueling employee engagement (Stewart et al., 2011). Thus, self-leadership is expected to be positively related to job productivity via a motivational pathway. Working with clearly defined goals and monitoring progress can energize employees and foster a deeper sense of purpose and involvement in the work (Breevaart et al., 2016). Accordingly, we propose the following, last hypothesis:

H3b: Remote work engagement mediates the relationship between self-leadership and job productivity.

According to the theoretical framework presented above, *Figure 1* shows the research model of this study.

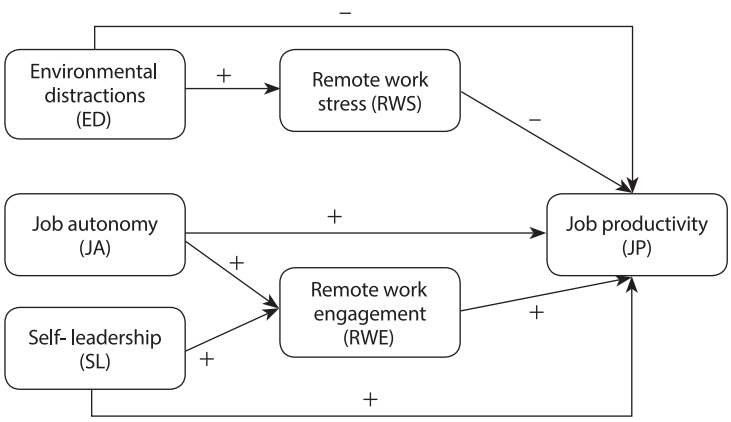

Figure 1. Conceptual research model with hypothesized associations.
Note: To improve readability, correlations between JA and SL, and between RWS and RWE are not shown here, although they were freely estimated in the model (see Figure 2)

Methods

Procedure

An online self-report questionnaire was administered through the Qualtrics platform between April and July 2020, during the large-scale transition to home-based work. Participants were recruited through social media and professional networks in Italy and Russia. To be eligible, respondents had to be adult employees, working from home full-time at the time of data collection, and able to complete either the Italian or Russian version of the questionnaire. Participation was voluntary, anonymous, and not compensated. Before accessing the questionnaire, respondents provided informed consent. The study was approved by the Ethics Committee of the university of the corresponding author. The study was conducted in accordance with the ethical standards for psychological and social research.

Participants

From a total of 578 participants, we retained 493 respondents who provided complete responses on the indicators included in the measurement and structural models. This is the final group of respondents used in all the subsequent analyses.

The study sample was composed of 247 Italian respondents and 246 Russian respondents. In the Italian subsample, 89 participants were men, 157 were women, and one identified as non-binary. In the Russian subsample, 72 participants were men and 174 were women. The Italian subsample was younger, with the largest age group being 26–35 years old (43.3%), followed by 36–45 years old (20.6%) and 46–55 years old (17.4%). In Russia, the largest age group was 36–45 years old (30.5%), followed by 46–55 years old (25.6%) and 26–35 years old (17.5%). Most participants worked in the private or public tertiary sector. In Italy, 118 participants worked in the private tertiary sector and 81 in the public tertiary sector; 44 worked in the secondary

sector and 3 in the primary sector, with one missing value on sector. In Russia, 114 participants worked in the private tertiary sector and 105 in the public tertiary sector; 25 worked in the secondary sector and 2 in the primary sector. Mean organizational tenure was 7.45 years in the Italian subsample ($SD = 9.40$) and 7.30 years in the Russian subsample ($SD = 6.76$). The sample was heterogeneous but not representative, as it was based on convenience recruitment during the emergency transition to home-based work.

Measures

All multi-item measures were based on previously validated instruments. Shortened versions of the scales were used to reduce participant burden and to capture the focal dimensions included in the research model. The English item pool was translated into Italian and Russian following a translation and back-translation procedure (Brislin, 1986). Discrepancies between the original and back-translated versions were discussed by the research team and resolved through consensus. An independent reviewer for each language then checked the final version for linguistic clarity and consistency with the intended construct meaning. Where necessary, item wording was adapted to address specifically working from home. The Russian item wordings are reported in the Appendix [the full Italian and Russian wordings are available in a Supplementary Appendix]. All measures were administered using a harmonized five-point agreement scale in both language versions, ranging from 1 = completely disagree to 5 = completely agree. When original instruments used a different response format, the item content was retained or adapted to the remote-work context, while the response scale was harmonized to maintain consistency across the questionnaire.

Environmental distractions (ED). Environmental distractions were measured with three items derived from Lee and Brand's (2005) work on perceived workspace conditions. The items captured acoustic distractions, lack of privacy, and visual distractions in the remote working area. Because the original items referred to workspace characteristics, their wording was contextualized to the home-based work setting (*e.g.*, "In my working area, I experience acoustic distractions"). Higher scores indicated greater perceived environmental distractions.

Job autonomy (JA). Job autonomy was measured with four items from Morgeson and Humphrey's (2006) Work Design Questionnaire. The items assessed the extent to which employees could make decisions about how to schedule, organize, and carry out their work activities. A sample item is: "My job allows me to make my own decisions about how to schedule my work". Higher scores indicated greater perceived job autonomy.

Self-leadership (SL). Self-leadership was assessed using a shorter set of items based on the Revised Self-Leadership Questionnaire (Houghton & Neck, 2002). The selected items (*e.g.*, "When I work, I always keep my tasks in mind") captured two self-regulatory strategies that were theoretically relevant to remote work: goal setting and self-monitoring. Thus, the measure should be interpreted as an indicator of the goal-oriented and self-monitoring aspects of self-leadership, rather than as a full assessment of the multidimensional self-leadership construct. Higher scores indicated greater self-leadership.

Remote work stress (RWS). Remote work stress was measured with four items from Weinert et al. (2015). The items assessed exhaustion, fatigue, and strain associated with working from home. A sample item is: "I feel exhausted from working from home". Higher scores indicated greater remote work stress.

Remote work engagement (RWE). Remote work engagement was measured with the three-item version of the Utrecht Work Engagement Scale (Schaufeli et al., 2019), with item wording referring specifically to working from home. The items assessed energy, enthusiasm, and immersion while working remotely. A sample item is: "When I work from home, I feel full of energy". Higher scores indicated greater remote work engagement.

Perceived job productivity (JP). Perceived job productivity was measured with a single item asking respondents to compare their productivity while working remotely with their usual productivity in the traditional office: "When I work remotely, I am more productive". The use of a single-item measure was based on prior evidence supporting the validity of single-item assessments of work-related outcomes, including job performance and job satisfaction (Ang & Eisend, 2018; Hartner-Tiefenthaler et al., 2023). This item has also been used in previous studies of remote work, showing meaningful associations with theoretically related constructs. For example, Galanti et al. (2021) reported correlations between perceived productivity and, respectively, work-family conflict ($r = -.40$) and job autonomy ($r = .18$). Similarly, Toscano and Zappalà (2020b) reported a positive association between perceived productivity and remote work satisfaction ($r = .65$). Nonetheless, while this measure offers practical advantages and evidence of nomological validity, it does not capture the full multidimensional nature of job productivity. Therefore, findings involving this outcome should be interpreted with caution.

Data Analysis

Before testing the research model, we examined the possibility of common method bias, using Harman's single-factor test, consisting of an exploratory factor analysis (EFA) with the principal axis method (*e.g.*, Zappalà et al., 2022). Because Harman's test is only a preliminary diagnostic, we also evaluated the structural distinctiveness of the study variables by comparing a six-factor confirmatory factor analysis (CFA) model with a one-factor model. This comparison was used to assess whether the measures reflected distinguishable constructs rather than a single undifferentiated response factor. Given the cross-national design, we then tested measurement invariance across the Italian and Russian subsamples. Reliability was tested by computing Cronbach's alpha and McDonald's omega, while convergent validity was examined by computing Average Variance Extracted values for each multi-item measure. After computing descriptive statistics and correlations, a structural equation modeling (SEM) was used to test the hypothesized model. The diagonally weighted least squares estimator (DWLS) was used in all SEM-based analyses of this study, because the focal indicators were measured with ordered five-point response categories and therefore needed to be treated as ordinal indicators, rather than on aggregated scale scores that may more closely approximate continuous variables. This decision was also supported by the size of the sample, with complete data on all the indicators used in this study, which was adequate for categorical least-squares estimation, and by departures

from normality in several indicators, such as autonomy_3 (skewness = −1.41, kurtosis = 1.82), goalsetting_1 (skewness = −1.45, kurtosis = 2.58), and selfmonitoring_2 (skewness = −1.62, kurtosis = 3.86). This choice is consistent with methodological and simulation evidence showing that categorical least-squares estimators are appropriate for ordinal indicators and can provide less biased and more accurate estimates under many ordinal-data conditions (Li, 2016; Rhemtulla et al., 2012).

Perceived job productivity was included as an observed single-item variable. For theoretical reasons (Bakker & Demerouti, 2007, 2017), the relationships between job autonomy and self-leadership, and between remote work stress and remote work engagement, were modeled as correlated. Finally, a multigroup SEM was conducted to explore whether the structural paths showed similar directions and magnitudes across the Italian and Russian subsamples. Given the cross-national design, measurement invariance across the Italian and Russian subsamples was examined through a sequence of increasingly constrained multigroup models. Configural, metric, and scalar/threshold invariance were evaluated. Changes in CFI not exceeding .010 were taken as evidence that the more constrained model did not substantially worsen model fit. All analyses were performed using Jamovi 2.3, integrating the semlj module for structural equation modeling (Gallucci & Jentschke, 2024), and were conducted on the sample.

Results

Validity of the Measurement Model and Reliability and Validity of the Measures

A single-factor EFA was performed as part of Harman's test to examine whether a dominant single-factor structure characterized the data for the variables involved. The single extracted factor accounted for 39% of the total variance, below the commonly used 50% threshold reported in prior research. This result does not suggest the presence of a common method factor, although it should be interpreted as a preliminary diagnostic rather than a definitive test of the absence of a common method bias. To test the structural independence of the six measures included in our model, we conducted two CFAs, comparing a one-factor model with a six-factor model. The hypothesized six-factor measurement model showed good fit to the data, $\chi^2(138) = 346$, $p < .001$, CFI = .995, TLI = .994, RMSEA = .055, SRMR = .056. To examine whether the study constructs were empirically distinguishable, we compared this model with a one-factor model in which all indicators loaded onto a single latent factor. The one-factor model showed poor fit, $\chi^2(152) = 6830$, $p < .001$, CFI = .841, TLI = .821, RMSEA = .299, SRMR = .238. These results supported the distinctiveness of the six constructs. *Table 1* reports reliability and convergent validity estimates for the multi-item measures. For JP, reliability could not be estimated directly because the construct was assessed with a single item. Its use was supported by prior evidence on single-item assessments of work-related outcomes (Ang & Eisend, 2018; Hartner-Tiefenthaler et al., 2023) and by previous remote work studies showing meaningful associations between this item and theoretically related constructs (Galanti et al., 2021; Toscano & Zappalà, 2020b).

Table 1

Reliability and Convergent Validity of the Adopted Multi-Item Measures

	Cronbach's α	**McDonald's ω**	**AVE**	**√AVE**
1. Environmental distractions	.76	.79	.63	.79
2. Job autonomy	.87	.89	.76	.87
3. Self-leadership	.77	.79	.60	.78
4. Remote work stress	.92	.94	.85	.92
5. Remote work engagement	.74	.77	.58	.76

Note: AVE = Average variance extracted

Given the cross-national design, measurement invariance across the Italian and Russian subsamples was examined through a sequence of increasingly constrained multigroup models focused on the measurement part of the model. The configural model showed good fit, $\chi^2(276) = 414$, $p < .001$, CFI = .997, TLI = .997, RMSEA = .045, SRMR = .062. Constraining factor loadings to equality across the Italian and Russian groups produced negligible deterioration in fit, ΔCFI = –.001, ΔRMSEA = .009, ΔSRMR = .001, supporting metric invariance. The scalar/threshold model also showed negligible change relative to the metric model, ΔCFI = –.001, ΔRMSEA = –.001, ΔSRMR = –.001. However, because the scalar/threshold model produced a technical warning, evidence for scalar/threshold invariance was interpreted cautiously.

Descriptive Statistics

Table 2 shows descriptive statistics and correlations among the study variables in our sample. All variables were significantly correlated in the expected direction, except for the association between self-leadership and remote work stress, which was not significant. The correlation between job autonomy and job productivity was positive and statistically significant, although small in magnitude.

Table 2

Descriptive Statistics and Correlations Among Study Variables

Variable	M	SD	1	2	3	4	5
1. Environmental distractions	2.50	1.16					
2. Job autonomy	4.02	.88	–.13**				
3. Self-leadership	4.19	.70	–.10*	.33**			
4. Remote work stress	2.69	1.29	.35**	–.10*	–.07		
5. Remote work engagement	3.24	.90	–.30**	.18**	.26**	–.42**	
6. Job productivity	3.18	1.22	–.27**	.09*	.19**	–.42**	.63**

*Note. N = 493; *p < .05, **p < .01*

Model Results

To test our hypotheses, we ran SEM on the full cross-national sample ($N = 493$). The tested model showed acceptable-to-good fit, with excellent incremental fit indices and acceptable absolute fit indices, $\chi^2(141) = 509$, $p < .001$, CFI = .991, TLI = .989, RMSEA = .073, SRMR = .070. Its results, shown in *Figure 2*, indicated a negative relationship between environmental distractions and job productivity (B = –.51; SE = .14; β = –.35; $p < .001$; H1a supported). Environmental distractions were also positively related to remote work stress (B = .50; SE = .06; β = .44; $p < .001$), while remote work stress was not significantly related to job productivity (B = .18; SE = .13; β = .14; p = .12). The indirect path between environmental distractions and job productivity through remote work stress was not significant (unstandardized point estimate = .09; SE = .07; standardized point estimate = .06; p = .12), which does not support H1b.

With regard to the relationships of job and personal resources with job productivity, the results showed that job autonomy was negatively and significantly associated with job productivity (B = –.19; SE = .07; β = –.14; p = .003; H2a not supported). The direct association between self-leadership and job productivity was negative, but only marginally significant (B = –.16; SE = .08; β = –.12; p = .051), and therefore H3a was not supported. In addition, both resources were positively related to remote work engagement: job autonomy was positively associated with remote work engagement (B = .15; SE = .05; β = .15; $p < .001$), and self-leadership was also positively associated with remote work engagement (B = .29; SE = .05; β = .30; $p < .001$). Remote work engagement, in turn, was strongly and positively associated with job productivity (B = 1.17; SE = .14; β = .81; $p < .001$).

Job autonomy showed a positive indirect association with job productivity through remote work engagement (unstandardized point estimate = .17; SE = .07; standardized point estimate = .12; $p < .001$), as did self-leadership (unstandardized point estimate = .34; SE = .08; standardized point estimate = .25; $p < .001$). These find-

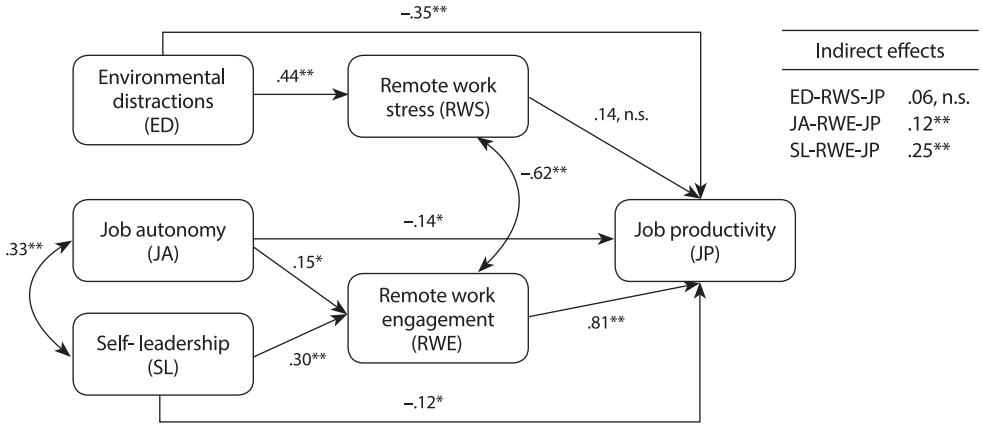

Figure 2. Tested model results

Note: n = 493. Standardized coefficients are reported for direct and indirect effects. Only latent factors are shown; observed indicators are omitted for readability

ings supported H2b and H3b, indicating that remote work engagement mediated the relationships of both job autonomy and self-leadership with perceived job productivity.

However, the pattern differed across the two resources. For job autonomy, the positive indirect effect through engagement was counterbalanced by a negative direct association with productivity, resulting in an overall association close to zero. For self-leadership, the positive indirect effect through engagement was robust, whereas the direct effect on productivity was weak and only marginally significant. This suggests that goal-setting and self-monitoring were positively related to productivity primarily through their association with remote work engagement rather than through a stable direct association with productivity.

Cross-National Multigroup Analysis

Given the cross-national design of the study, we examined whether the structural paths observed in the full sample showed similar directions and magnitudes across Italian and Russian participants. This analysis was exploratory and focused on the similarity of structural relationships rather than on mean differences. Although scalar/threshold invariance was acceptable according to approximate fit criteria, the multigroup structural results were interpreted cautiously because the scalar/threshold

Table 3

Standardized Structural Paths in the Italian and Russian Subsamples

Path	**Direct effects**	
	Italy	**Russia**
ED-JP	-.34**	-.30**
ED-RWS	.37**	.49**
RWS-JP	.21	.06
JA-JP	-.14**	-.14
JA-RWE	.08	.22**
SL-JP	-.12	-.07
SL-RWE	.31**	.30**
RWE-JP	.83**	.75**
	Indirect effects	
	Italy	Russia
ED-RWS-JP	.08	.03
JA-RWE-JP	.07*	.17**
SL-RWE-JP	.25**	.22**

*Note: Country subsample sizes: n = 247 for Italy and n = 246 for Russia.*p < .05, **p < .01; ED = Environmental distractions, JA = Job autonomy, SL = Self-leadership, RWS = Remote work stress, RWE = Remote work engagement, JP= Job productivity*

model produced a technical warning and because not all full-sample paths reached statistical significance in both national subsamples. The multigroup model showed acceptable-to-good fit, $\chi^2(282) = 583$, $p < .001$, CFI = .994, TLI = .993, RMSEA = .066, SRMR = .076. The relationships observed in this multigroup comparison are reported in *Table 3*.

Table 3 suggests partial convergence between the country-specific models and the full-sample model. Several coefficients showed similar directions in Italy and Russia, especially the paths from environmental distractions to remote work stress and job productivity, from self-leadership to remote work engagement, and from remote work engagement to job productivity. However, not all significant paths observed in the full sample reached statistical significance in both subsamples. Hence, the multigroup findings should be interpreted as exploratory evidence of partial structural convergence rather than as full replication across countries.

The negative association between environmental distractions and job productivity, and the positive association between environmental distractions and remote work stress, were observed in both subsamples. As in the full-sample model, remote work stress was not significantly related to job productivity. The negative relationship between job autonomy and job productivity was similar in magnitude across the two subsamples, although it reached statistical significance only in the Italian subsample. Conversely, the positive relationship between job autonomy and remote work engagement was statistically significant in both subsamples, although it was small in Italy and stronger in Russia.

In the country-specific models, the direct relationship between self-leadership and job productivity was not statistically significant in either subsample. This is consistent with the cautious interpretation of the full-sample direct effect, which was negative but only marginally significant. However, in both national subsamples, the indirect relationship between self-leadership and job productivity through remote work engagement was positive and statistically significant. Overall, the multigroup analysis suggests that some central pathways were directionally similar in Italy and Russia, especially the engagement-based motivational pathway, but the evidence supports only partial convergence rather than broad structural comparability.

Discussion

This study examined working from home as a psychological and organizational reconfiguration of that work activity. Drawing on the JD-R model, we tested whether environmental distractions, job autonomy, and self-leadership were associated with perceived productivity through remote work stress and remote work engagement. We also explored whether these structural associations showed similar patterns across Italian and Russian employees.

Overall, the findings highlight that environmental distractions were associated with higher remote work stress and lower perceived job productivity, although the indirect effect through remote work stress was not significant. Job autonomy and self-leadership were positively associated with remote work engagement, which in turn was strongly related to perceived job productivity. A theoretically relevant finding concerns the ambivalent role of job autonomy and self-leadership. Both resources

showed positive indirect associations with productivity through remote work engagement. However, contrary to our expectations, job autonomy showed a statistically significant negative direct association with productivity, whereas the negative direct association of self-leadership was marginal and should therefore be interpreted more cautiously. These direct coefficients should not be interpreted as simple bivariate relationships. In the zero-order correlations, both job autonomy and self-leadership were positively related to productivity; the negative direct coefficients emerged only after remote work engagement and the other predictors were included in the model. This pattern is compatible with an inconsistent mediation or suppression-like structure, in which autonomy and self-leadership support productivity indirectly through engagement, while their residual direct associations may capture the additional self-regulatory burden of highly autonomous remote work. During the emergency transition to home-based work, employees often had to plan, prioritize, monitor, coordinate, and protect their work boundaries with limited managerial support. Under these conditions, autonomy may have functioned as an ambivalent resource: motivationally beneficial, but also cognitively and organizationally demanding. For self-leadership, however, the evidence points more clearly to an indirect motivational role through engagement than to a stable direct association with productivity. This interpretation is consistent with refinements of the JD-R model suggesting that resources can have context-dependent effects and with the idea that excessive or poorly matched resources may undermine rather than support functioning (Demerouti, 2025; Kubicek et al., 2017; Zhou, 2020).

At the same time, the mediation analysis showed that remote work engagement significantly mediated the relationships of both job autonomy and self-leadership with job productivity, supporting the motivational pathway of the JD-R model and suggesting that engagement remains a key proximal mechanism linking resources to performance-related outcomes in remote work. This pattern was especially clear for self-leadership: goal setting and self-monitoring were positively related to productivity primarily because they were associated with higher remote work engagement. This result suggests that self-regulatory strategies may be especially relevant when employees work with reduced direct supervision and must actively organize their own work activity. Finally, the multigroup analysis extends previous evidence on Russian remote workers (Toscano et al., 2022) by suggesting partial convergence in some central psychosocial mechanisms linking remote work conditions to productivity across Italian and Russian employees. However, these findings should be interpreted as exploratory evidence of partial structural similarity, not as evidence of broad cross-national equivalence.

Theoretical and Practical Implications

Overall, the findings have several theoretical implications. First, they highlight the underexplored role of environmental distractions as a relevant job demand in work-from-home settings. By showing their negative association with perceived productivity, the study extends recent research suggesting that future studies should distinguish more clearly between digital and non-digital distractions (Nakayama & Chen, 2023).

Contrary to expectations, remote work stress was not related to productivity. We did not observe the hypothesized indirect effect of environmental distractions through stress. One possible explanation is that distractions impair productivity directly by reducing concentration and increasing task fragmentation rather than through a general stress mechanism. Alternatively, the stress measure may have captured fatigue and exhaustion but not interruption-related cognitive costs, such as attentional switching and perceived inefficiency. Future research should therefore distinguish more clearly between stress, cognitive load, interruptions, and concentration when examining health-impairment processes in remote work.

Second, the study provides insights into the motivational role of autonomy and self-leadership. Consistent with JD-R theory, both resources were positively associated with engagement, which in turn predicted productivity. However, the direct association between autonomy and productivity was negative, whereas the direct association between self-leadership and productivity was weak and marginal. This pattern highlights the context-dependent nature of resources and supports recent arguments that resources do not always produce uniformly positive effects (Demerouti, 2025).

Finally, the use of Italian and Russian samples allowed a preliminary examination of the cross-national robustness of the proposed JD-R mechanisms. The partially similar pattern of associations suggests that some JD-R mechanisms in remote work may operate in comparable directions across different national contexts. However, the evidence does not support strong claims of structural equivalence. The multigroup findings should therefore be treated as preliminary and hypothesis-generating.

Our findings extend research on remote work and offer implications for employees, managers, and HR professionals. First, the negative association between environmental distractions and productivity suggests that the home workspace should not be viewed solely as a private matter. Organizations can support concentration by providing guidance on workspace setup, reducing interruptions, and establishing clear norms regarding availability, response times, and meetings. Remote work design should therefore address not only technological access, but also conditions that protect focused work and cognitive well-being.

Second, while self-leadership appears important for maintaining direction and engagement in remote work, it should not be considered a substitute for organizational support. Goal setting and self-monitoring can be strengthened through training, clear performance expectations, regular feedback, and supportive management practices. Likewise, autonomy should be accompanied by sufficient structure, because discretion without clear priorities, feedback, and boundary management support may increase the burden of self-organization.

Overall, effective remote work requires a balanced design in which employee autonomy and self-regulatory capacities are supported by organizational structure, guidance, and protection from avoidable distractions.

Limitations

Like all studies, this research has limitations. First, the cross-sectional design reduces the strength of our findings and is not fully consistent with the directional assumptions implied by mediation models. Accordingly, the indirect effects should be inter-

preted as associations among theoretically ordered variables rather than as evidence of causal processes.

Second, the data were collected in 2020, during an emergency transition to home-based work. The results should therefore be generalized cautiously and should be retested in contemporary remote and hybrid work arrangements that are more planned, institutionalized, and supported by organizational policies.

Third, although the study included Italian and Russian employees, national context was used as a grouping variable rather than as a direct measure of cultural, institutional, or organizational characteristics. Therefore, the study cannot identify the specific cultural, managerial, or policy-related mechanisms responsible for possible country-specific differences. Future studies should include direct measures of cultural values, remote work policies, managerial practices, and organizational support to explain why remote work mechanisms may vary across contexts.

Fourth, although measurement invariance was generally supported according to approximate fit criteria, evidence for scalar/threshold invariance was interpreted cautiously because the scalar/threshold model produced a technical warning. The cross-national findings should also be interpreted cautiously because the country-specific structural models were based on relatively modest subsample sizes and not all paths replicated across both groups.

Finally, perceived job productivity was measured with a single item. Although this item has practical advantages and previous evidence of nomological validity, it does not capture the full multidimensional nature of productivity. Future studies should combine self-reported productivity with supervisor ratings, objective indicators, or task-specific performance measures.

Several procedural remedies were adopted to reduce common method bias, including anonymity, voluntary participation, and the separation of measures within the questionnaire. In addition, Harman's single-factor test and the comparison between the one-factor and six-factor CFA models suggested that the study variables were not reducible to a single undifferentiated response factor. Nevertheless, because all variables were collected through self-report measures, common method bias cannot be completely ruled out.

Conclusion

This study provided insights into the dynamics of remote work by examining the roles of environmental distractions, job autonomy, and self-leadership as predictors of perceived job productivity, with a particular focus on remote work stress and remote work engagement as mediating mechanisms. Drawing on a cross-national sample from Italy and Russia, the findings support the motivational pathway of the JD-R model: both autonomy and self-leadership were positively associated with engagement, which in turn predicted productivity. At the same time, the findings were less consistent with the health-impairment pathway, because remote work stress did not significantly mediate the association between environmental distractions and productivity. The unexpected negative direct association between job autonomy and productivity, together with the weak and marginal negative direct association between self-leadership and productivity, suggests the need to critically reassess the

assumption that resources are always uniformly beneficial in remote work contexts. Under conditions of high self-management and limited structure, autonomy may also increase the self-regulatory burden placed on employees, whereas self-leadership appears to contribute to productivity primarily by sustaining engagement rather than through a stable direct effect on productivity.

The study provides preliminary evidence of partial convergence in some key psychological mechanisms across the Italian and Russian subsamples, while also indicating that cross-national findings should be interpreted cautiously because not all structural paths were consistently significant across both countries. Future research should retest these mechanisms in contemporary hybrid work arrangements, where remote work is more planned, institutionalized, and supported by organizational policies than during the emergency pandemic transition. Such research should also consider sustainable work design, examining how flexible work can support productivity while protecting employee health, work-life boundaries, inclusion, and long-term organizational functioning. Overall, the findings contribute to the theoretical refinement of the JD-R model in remote work settings and offer guidance for organizations seeking to sustain engagement and productivity in evolving work arrangements

Ethics Statement

This study obtained ethics approval from the University of Bologna, prot. N. 252554.

Author Contributions

F.T. and S.Z. conceived and designed the study, developed its theoretical framework, and coordinated the research and data collection in Italy. M.V.P. and E.V.K. contributed to the conceptualization of the study, adapted the questionnaires to the Russian context, and coordinated data collection in Russia. F.T. performed the statistical analyses, with input from S.Z., F.T. and S.Z. interpreted the results, and drafted and revised the manuscript. All authors read and approved the final version of the manuscript.

Informed Consent from the Participants

All the participants filled in the informed consent form before taking part.

Conflict of Interest

The authors declare no conflict of interest.

Acknowledgements

The authors thank Fatimah Anwar Ali for her preliminary work on earlier versions of portions of this study as part of her master's thesis. Her contribution was limited to those preliminary drafts and did not meet the journal's criteria for authorship, as documented through the CRediT contributor-role taxonomy.

References

Allen, T. D., Golden, T. D., & Shockley, K. M. (2015). How effective is telecommuting? Assessing the status of our scientific findings. *Psychological Science in the Public Interest*, *16*(2), 40–68. https://doi.org/10.1177/1529100615593273

Ang, L., & Eisend, M. (2018). Single versus multiple measurement of attitudes. *Journal of Advertising Research*, *58*(2), 218–227. https://doi.org/10.2501/JAR-2017-001

Baethge, A., & Rigotti, T. (2013). Interruptions to workflow: Their relationship with irritation and satisfaction with performance, and the mediating roles of time pressure and mental demands. *Work and Stress*, *27*(1), 43–63. https://doi.org/10.1080/02678373.2013.761783

Bakker, A. B., & Demerouti, E. (2007). The Job Demands-Resources model: State of the art. *Journal of Managerial Psychology*, *22*(3), 309–328. https://doi.org/10.1108/02683940710733115

Bakker, A. B., & Demerouti, E. (2017). Job demands-resources theory: Taking stock and looking forward. *Journal of Occupational Health Psychology*, *22*(3), 273–285. https://doi.org/10.1037/ocp0000056

Bakker, A. B., & Demerouti, E. (2024). Job demands–resources theory: Frequently asked questions. *Journal of Occupational Health Psychology*, *29*(3), 188–200. https://doi.org/10.1037/ocp0000376

Bedny, G. Z., & Karwowski, W. (2006). *A systemic-structural theory of activity: Applications to human performance and work design*. CRC Press.

Bedny, G. Z., & Meister, D. (1997). *The Russian theory of activity: Current applications to design and learning*. Lawrence Erlbaum Associates.

Breevaart, K., Bakker, A. B., Demerouti, E., & Derks, D. (2016). Who takes the lead? A multi-source diary study on leadership, work engagement, and job performance. *Journal of Organizational Behavior*, *37*(3), 309–325. https://doi.org/10.1002/job.2041

Bloom, N., Han, R., & Liang, J. (2024). Hybrid working from home improves retention without damaging performance. *Nature, 630*, 920–925. https://doi.org/10.1038/s41586-024-07500-2

Brislin, R. W. (1986). The wording and translation of research instruments. In L. W. Lonner & J. W. Berry (Eds.), *Field methods in cross-cultural psychology* (pp. 137–164). SAGE Publications Inc.

Deci, E. L., Olafsen, A. H., & Ryan, R. M. (2017). Self-Determination Theory in work organizations: The state of a science. *Annual Review of Organizational Psychology and Organizational Behavior*, *4*(1), 19–43. https://doi.org/10.1146/annurev-orgpsych-032516-113108

Demerouti, E. (2025). Job demands-resources and conservation of resources theories: How do they help to explain employee well-being and future job design? *Journal of Business Research*, *192*(February), 115296. https://doi.org/10.1016/j.jbusres.2025.115296

Dong, J. J., Tan, Z. D., Zhang, Y. L., Sun, Y. J., & Huang, Y. K. (2025). Work from home and employee well-being: A double-edged sword. *BMC Psychology, 13*, 748. https://doi.org/10.1186/s40359-025-02994-5

Galanti, T., Guidetti, G., Mazzei, E., Zappalà, S., & Toscano, F. (2021). Work from home during the COVID-19 outbreak: The impact on employees' remote work productivity, engagement, and stress. *Journal of Occupational and Environmental Medicine*, *63*(7), e426–e432. https://doi.org/10.1097/JOM.0000000000002236

Gallucci, M., & Jentschke, S. (2024). *SEMLj* (1.1.6). https://semlj.github.io

Hartner-Tiefenthaler, M., Mostafa, A. M. S., & Koeszegi, S. T. (2023). The double-edged sword of online access to work tools outside work: The relationship with flexible working, work interrupting nonwork behaviors and job satisfaction. *Frontiers in Public Health*, *10*(1), 1035989. https://doi.org/10.3389/fpubh.2022.1035989

Houghton, J. D., & Neck, C. P. (2002). The revised self-leadership questionnaire. *Journal of Managerial Psychology*, *17*(8), 672–691. https://doi.org/10.1108/02683940210450484

Keller, A. C., Meier, L. L., Elfering, A., & Semmer, N. K. (2020). Please wait until I am done! Longitudinal effects of work interruptions on employee well-being. *Work and Stress*, *34*(2), 148–167. https://doi.org/10.1080/02678373.2019.1579266

Kubicek, B., Paškvan, M., & Bunner, J. (2017). The bright and dark sides of job autonomy. In *Job demands in a changing world of work* (pp. 45–63). Springer International Publishing. https://doi.org/10.1007/978-3-319-54678-0_4

Lee, S. Y., & Brand, J. L. (2005). Effects of control over office workspace on perceptions of the work environment and work outcomes. *Journal of Environmental Psychology*, *25*(3), 323–333. https://doi.org/10.1016/j.jenvp.2005.08.001

Leonova, A. B., Kuznetsova, A. S., & Barabanshchikova, V. V. (2010). Self-regulation training and prevention of negative human functional states at work: Traditions and recent issues in Russian applied research. *Psychology in Russia: State of the Art*, *3*(1), 482–507. https://doi.org/10.11621/pir.2010.0023

Li, C. H. (2016). Confirmatory factor analysis with ordinal data: Comparing robust maximum likelihood and diagonally weighted least squares. *Behavior Research Methods*, *48*(3), 936–949. https://doi.org/10.3758/s13428-015-0619-7

Locke, E. A., & Latham, G. P. (2013). New developments in goal setting and task performance. In *New developments in goal setting and task performance*. Routledge. https://doi.org/10.4324/9780203082744

Maidansky, A. D. (2021). Controversy and growth points in the activity theory in psychology. *Psychology in Russia*, *14*(4), 3-17. https://doi.org/10.11621/pir.2021.0401

Memon, M. A., Salleh, R., Mirza, M. Z., Cheah, J.-H., Ting, H., Ahmad, M. S., & Tariq, A. (2021). Satisfaction matters: The relationships between HRM practices, work engagement and turnover intention. *International Journal of Manpower*, *42*(1), 21–50. https://doi.org/10.1108/IJM-04-2018-0127

Monteiro, N. P., Straume, O. R., & Valente, M. (2021). When does remote electronic access (not) boost productivity? Longitudinal evidence from Portugal. *Information Economics and Policy*, *56*(February). https://doi.org/10.1016/j.infoecopol.2021.100923

Morgeson, F. P., & Humphrey, S. E. (2006). The Work Design Questionnaire (WDQ): Developing and validating a comprehensive measure for assessing job design and the nature of work. *Journal of Applied Psychology*, *91*(6), 1321–1339. https://doi.org/10.1037/0021-9010.91.6.1321

Moser, K., & Galais, N. (2007). Self-monitoring and job performance: The moderating role of tenure. In *International Journal of Selection and Assessment*, *15*(1), pp. 83–93. https://doi.org/10.1111/j.1468-2389.2007.00370.x

Murphy, K. R. (1990). Job performance and productivity. In *Psychology in organizations* (pp. 157–177). Lawrence Erlbaum Associates Publishers.

Nakayama, M., & Chen, C. C. (2023). Digital distraction, non-digital distraction and psychological bearing of remote workers. *Interacting with Computers*, *35*(6), 789–800. https://doi.org/10.1093/iwc/iwad051

Rhemtulla, M., Brosseau-Liard, P. É., & Savalei, V. (2012). When can categorical variables be treated as continuous? A comparison of robust continuous and categorical SEM estimation methods under suboptimal conditions. *Psychological Methods*, *17*(3), 354–373. https://doi.org/10.1037/a0029315

Saragih, S. (2011). The effects of job autonomy on work outcomes: Self efficacy as an intervening variable. *International Research Journal of Business Studies*, *4*(3), 203–215. https://doi.org/10.21632/irjbs.4.3.203-215

Schaufeli, W. B., Shimazu, A., Hakanen, J., Salanova, M., & De Witte, H. (2019). An ultra-short measure for work engagement. *European Journal of Psychological Assessment*, *35*(4), 577–591. https://doi.org/10.1027/1015-5759/a000430

Sonnentag, S. (2017). A task-level perspective on work engagement: A new approach that helps to differentiate the concepts of engagement and burnout. *Burnout Research*, *5*(April), 12–20. https://doi.org/10.1016/j.burn.2017.04.001

Stewart, G. L., Courtright, S. H., & Manz, C. C. (2011). Self-leadership: A multilevel review. *Journal of Management*, *37*(1), 185–222. https://doi.org/10.1177/0149206310383911

Toscano, F., Bigliardi, E., Polevaya, M. V., Kamneva, E. V., & Zappalà, S. (2022). Working remotely during the COVID-19 pandemic: Work-related psychosocial factors, work satisfaction, and job performance among Russian employees. *Psychology in Russia: State of the Art, 15*(1), 3–19. https://doi.org/10.11621/pir.2022.0101

Toscano, F., & Zappalà, S. (2020a). Smart working in Italia: Origine, diffusione e possibili esiti. *Psicologia Sociale*, *15*(2), 203–223. https://doi.org/10.1482/96843

Toscano, F., & Zappalà, S. (2020b). Social isolation and stress as predictors of productivity perception and remote work satisfaction during the COVID-19 pandemic: The role of concern about the virus in a moderated double mediation. *Sustainability*, *12*(23), 9804. https://doi.org/10.3390/su12239804

Toscano, F., & Zappalà, S. (2021). Overall job performance, remote work engagement, living with children, and remote work productivity during the COVID-19 pandemic: A mediated moderation model. *European Journal of Psychology Open*, *80*(3), 133–142. https://doi.org/10.1024/2673-8627/a000015

Troll, E. S., Venz, L., Weitzenegger, F., & Loschelder, D. D. (2022). Working from home during the COVID-19 crisis: How self-control strategies elucidate employees' job performance. *Applied Psychology*, *71*(3), 853–880. https://doi.org/10.1111/apps.12352

Weinert, C., Maier, C., & Laumer, S. (2015). Why are teleworkers stressed? An empirical analysis of the causes of telework-enabled stress. *Proceedings Der 12. Internationalen Tagung Wirtschaftsinformatik*, 1407–1421.

Wineman, J. D., & Barnes, J. (2018). Workplace settings. In *Environmental psychology and human well-being* (pp. 167–192). Elsevier. https://doi.org/10.1016/B978-0-12-811481-0.00007-X

Xiao, Y., Becerik-Gerber, B., Lucas, G., & Roll, S. C. (2021). Impacts of working from home during COVID-19 pandemic on physical and mental well-being of office workstation users. *Journal of Occupational & Environmental Medicine*, *63*(3), 181–190. https://doi.org/10.1097/JOM.0000000000002097

Wells, J., Scheibein, F., Pais, L., Rebelo dos Santos, N., Dalluege, C.-A., Czakert, J. P., & Berger, R. (2023). A systematic review of the impact of remote working referenced to the concept of work–life flow on physical and psychological health. *Workplace Health & Safety, 71*(11), 507–521. https://doi.org/10.1177/21650799231176397

Zappalà, S., Swanzy, E. K., & Toscano, F. (2022). Workload and mental well-being of homeworkers: The mediating effects of work-family conflict, sleeping problems, and work engagement. *Journal of Occupational and Environmental Medicine*, *64*(10), E647–E655. https://doi.org/10.1097/JOM.0000000000002659

Zhou, E. (2020). The "Too-Much-of-a-Good-Thing" effect of job autonomy and its explanation mechanism. *Psychology*, *11*(02), 299–313. https://doi.org/10.4236/psych.2020.112019

Original manuscript received June 9, 2026
Revised manuscript accepted July 29, 2026
First published online September 1, 2026

To cite this article: Toscano, F., Zappalà, S., Polevaya, M.V., Kamneva, E.V. (2026). Working from Home Across National Contexts: Demands, Resources, Engagement, and Productivity Among Italian and Russian Employees. *Psychology in Russia: State of the Art*, *19*(3), 25-44. DOI: 10.11621/pir.2026.0302

Psychology in Russia: State of the Art
Volume 19, Issue 3, 2026

Russian Psychological Society
Lomonosov Moscow State University

DEVELOPMENTAL PSYCHOLOGY

Teacher Implementation of Sociodramatic Play in Early Childhood Education: A Systematic Review

Vera L. Sukhikh[a,b*], Nikolay N. Veresov[c], Margarita N. Gavrilova[a,b]

[a] *Lomonosov Moscow State University, Russia*
[b] *Federal Scientific Center for Psychological and Interdisciplinary Research, Moscow, Russia*
[c] *Monash University, Melbourne, Australia*

*Corresponding author. E-mail: sukhikhvera@gmail.com

Background. Sociodramatic play is widely recognized as a leading activity in early childhood development, yet a persistent gap exists between its theoretical value and systematic implementation in educational practice. In this review we use the term as an umbrella construct for educator-supported collective pretend play defined by four criteria: (1) enactment of explicit roles, (2) interaction between roles rather than solitary pretense, (3) an unfolding narrative storyline, and (4) collective participation sustaining a joint imaginary situation.

Objective. To identify pedagogical approaches to implementing sociodramatic play, to map required teacher competencies, to synthesize implementation barriers, and to document outcomes for children and educators.

Design. Following PRISMA guidelines, searches were conducted across ScienceDirect and PubMed, supplemented by hand-searching and expert consultation. Studies were eligible if they involved educator-led collective sociodramatic play with children aged 3–7 in Early Childhood Education (ECE) settings. Narrative synthesis was conducted across 27 studies (35 papers) from 10 countries.

Results. Five implementation approaches were identified and systematically characterized. Child outcomes across content acquisition, developmental progress, and play maturity were broadly positive across diverse international contexts, despite null findings in some studies. A multidimensional teacher competency profile was identified and implementation barriers were multilevel: personal, pedagogical, and systemic.

Conclusion. Effective implementation requires performative and reflective competencies distinct from conventional teaching skills, which may explain persistent low confidence among educators. Widespread adoption demands pedagogically accessible approaches supported by sustained professional development and systemic infrastructure.

Keywords: sociodramatic play, early childhood education, play-based learning, teacher competences, implementation barriers, pedagogical approaches, child development outcomes

© Sukhikh, V.L., Veresov, N.N., Gavrilova, M.N., 2026

ISSN 2074-6857 (Print) / ISSN 2307-2202 (Online)
http://psychologyinrussia.com

Introduction

In recent decades, a large body of theoretical and empirical research justified play as a cornerstone of child development and education. Despite widespread recognition of play's developmental value, collective sociodramatic play seems to be a challenging practice for educators.

Central to this activity is collaborative narrative construction: sociodramatic play is a child-initiated, collective activity in which children spontaneously assume roles, negotiate imaginary scenarios, and generate and transform storylines in real time through shared symbolic action. In this sense, storytelling is not a separate activity that accompanies play, but a constitutive dimension of it. This distinguishes sociodramatic play from adjacent literacy-oriented practices such as the reading, retelling, or structured enactment of a pre-existing text in which the narrative is largely given rather than generated, and the emphasis falls on linguistic form, sequencing, and comprehension rather than on the collective negotiation of shared imaginary meaning. This distinction matters because these two orientations place different demands on teachers, and because some pedagogical models deliberately combine them (see below).

The overwhelming majority of interventions demonstrating the positive impact of sociodramatic play have been conducted by the researchers themselves under special conditions outside usual educational routines, aligning with previous meta-analyses that have indicated that researcher-implemented interventions often lead to larger gains than those conducted by teachers (Marulis & Neuman, 2010). ECE teachers encounter difficulties in implementing this type of play within the curriculum. The diversity of approaches to sociodramatic play necessitates a thorough analysis and evaluation of their effectiveness and implementation in preschool settings.

The present study aims to address these gaps by identifying effective pedagogical approaches, required teacher competencies, and implementation barriers that influence outcomes for both children and educators.

Sociodramatic Play Among Other Types of Play

The literature often uses several overlapping terms interchangeably, so clear boundaries are needed. Pretend play is the broadest category. It encompasses any activity involving symbolic substitution (Ryabkova et al., 2025; Lillard et al., 2013; Weisberg et al., 2013). Role play is narrower and denotes the enactment of a role, but not necessarily in interaction with others. Dramatic play refers to imitative enactment of a role, which may remain solitary. Sociodramatic play, as defined by Smilansky and Shefatya (1990), adds two requirements: social interaction and verbal communication among at least two players who coordinate their roles around a common theme.

The cultural-historical approach offers a converging definition. Vygotsky (2016) characterized such play through the creation of an imaginary situation, the adoption of roles, and adherence to the rules implicit in those roles. He positioned it as a lead-

ing activity of the preschool years, through which higher psychological functions develop. In play, children reproduce and model the social roles and relationships they observe around them, learn to understand social norms and rules, develop imagination, self-regulation, and the ability to think abstractly. Bodrova and Leong (2007), using the term make-believe play, specified that its mature forms involve multi-theme scenarios, elaborate role speech, and sustained collaborative engagement.

A recurring feature across these definitions is the role of shared narrative. Storytelling and narrative are therefore best understood as internal components of sociodramatic play rather than as a distinct object of inquiry parallel to it. This is important for the present review because pedagogical models vary in how explicitly they foreground narrative construction: some (STSA, Playworlds) make the storyline an explicit shared artefact, whereas others embed narrative implicitly within collective role enactment. All, however, share the same underlying construct.

Drawing on these definitions, we adopt four criteria to distinguish sociodramatic play from adjacent constructs: (1) children adopt and enact explicit roles; (2) players interact between roles, rather than engaging in solitary pretense; (3) play unfolds as a developing storyline; and (4) play is collective, involving two or more participants who jointly sustain the imaginary situation. We use "sociodramatic play" as an umbrella term for educator-supported collective pretend play that meets these four criteria, encompassing a range of pedagogical models that differ in structure and teacher involvement, including those in which storytelling is made an explicit organizing element.

Sociodramatic Play and Play-Based Learning

Play-based learning encompasses a broad spectrum of approaches in early childhood education, yet not all implementations center on sociodramatic play. Many educational programs emphasize other play types, such as guided games with rules for mathematical development (DePascale & Ramani, 2024), constructive play for spatial reasoning (Bower et al., 2025), or exploratory play for scientific inquiry (Bulunuz, 2013). Weisberg et al. (2013) describe "guided play" approaches that incorporate adult-directed learning goals within playful contexts but often lack the characteristic of mature role-play such as complex social negotiation and sustained imaginative scenarios. Notable exceptions include Pramling Samuelsson's developmental pedagogy research in Sweden (Pramling Samuelsson & Asplund Carlsson, 2008) and the Tools of the Mind curriculum, which deliberately structures sociodramatic play opportunities to develop self-regulation and executive functions through Vygotskian principles of mature play (Bodrova & Leong, 2025; Diamond et al., 2007). And while numerous practitioner resources offer guideline for supporting children's play (*e.g.*, Bredikyte, M., 2017; Smilansky & Shefatya, 1990), the extent to which the recommendations can be meaningfully implemented within standard educational routines and curricula remains questionable.

Existing Reviews

Several reviews and meta-analyses have examined the effects of pretend play and play-based pedagogies on child development, but none focus specifically on teacher implementation of sociodramatic play. Smits-van den Natt et al. (2024) synthesized evidence on pretend play and social competence, highlighting improvements in social skills, yet their review did not consider teacher competencies or classroom implementation. Skene et al. (2022) reviewed guidance during play in educational contexts, emphasizing the benefits of adult scaffolding for child learning outcomes, but the studies reviewed did not focus on collective sociodramatic play specifically. Pyle et al. (2017) conducted a scoping review of play-based pedagogies in kindergarten, summarizing pedagogical approaches and child outcomes, but teacher professional development and implementation strategies were only briefly discussed and again there was no focus on collective sociodramatic play specifically. Lillard et al. (2013) provided a comprehensive review of pretend play's impact on child development across cognitive, social, and emotional domains, yet teacher involvement in implementing narrative story-based play was not a focus. These reviews indicate some evidence for the benefits of play-based approaches, but there is a clear gap regarding systematic synthesis of teacher practices, skills, and competencies required to implement sociodramatic play in preschool classrooms.

Rationale and Objectives

Given the recognized role of sociodramatic play as a leading activity that fosters the development of higher psychological functions in the preschool years, this systematic review aims to synthesize available evidence on the implementation of sociodramatic play in preschool settings to answer the main research question: How is educator-led collective sociodramatic play implemented in early childhood education settings, and with what outcomes for children and educators?

To answer this question, the review pursues the following specific objectives:

1. To identify and describe approaches to implement sociodramatic play within early childhood educational settings;
2. To determine the skills, training, and competencies required for educators to effectively integrate sociodramatic play into the preschool curriculum;
3. To identify and synthesize evidence on the barriers, challenges, and limitations encountered by educators when integrating sociodramatic play;
4. To map the outcomes associated with the use of sociodramatic play in preschool settings.

Methods

A protocol was submitted to the PROSPERO registry for systematic reviews (CRD number: 420251122410). PRISMA guidelines of reporting were followed (Page et al., 2021).

Eligibility Criteria

Types of Participants

The population of interest was early childhood teachers, including both pre-service and in-service teachers, and children in pre-school/kindergarten from the age 3 to 7. Studies focusing on children outside of this age range were excluded. Studies focusing only on children with special educational needs were excluded. Studies were excluded if the sociodramatic play intervention was delivered by a visiting facilitator or researcher, rather than by a regular educator working within the educational setting. The review focuses specifically on educator-led implementation of play into the curriculum.

Types of Studies

The searches among English-language published journal articles were conducted with no restrictions based on study design or publication date. RCTs, qualitative, quantitative, or mixed methods were eligible for inclusion. Studies focusing on the implementation of play-based curriculum at the national level were excluded, as the articles did not contain specific information related to sociodramatic play. Systematic reviews and meta-analyses as well as exploratory research, where teachers were not assigned explicit tasks, were excluded. Tool of the Mind studies were excluded because they primarily examine broad curriculum effects.

Types of Settings

Studies conducted in early childhood education settings (*e.g.*, preschools, kindergartens, nursery schools) were included. Studies conducted in clinical, home-based, or informal settings without a clear educational context were excluded.

Types of Interventions

This review included studies that evaluate sociodramatic play conducted by teachers in ECE as an addition to the main curricula. Eligible interventions had to be implemented with the intention of supporting children's learning, development, or play skills within structured or semi-structured educational settings.

Consistent with the definition adopted above, the review included only forms of play that meet four criteria: (1) children adopt and enact explicit roles; (2) players interact between roles, rather than engage in solitary pretense; (3) play unfolds as a developing storyline; and (4) play is collective, involving two or more participants who jointly sustain the imaginary situation.

Studies were excluded if they limited play activity to non-narrative, non-interactive formats, such as circle games, group movement or serious games, or songs with gestures, where children do not engage in sustained role-taking or narrative development. Studies were also excluded if they solely focused on free play and the educator's role was limited to preparation of the environment.

Types of Outcome Measures

Included studies reported on at least one of the following outcome domains:

- Academic skills (*e.g.*, language, early literacy, math);
- Developmental outcomes (*e.g.*, cognitive, physical, social, emotional development);
- Play development (*e.g.*, complexity of play, role enactment);
- Educator-related factors and outcomes (*e.g.*, pedagogical strategies, training needs, competencies);
- Barriers and challenges to integration of sociodramatic play in educational contexts.

Search Methods for Identification of Studies

Two widely recognized and representative databases – ScienceDirect (Scopus) and PubMed — were selected for the literature search, as they are commonly used in education research.

An initial scope of the literature and consultation with an expert panel identified the following keywords: role-playing game; sociodramatic play; dramatic play; pretend play; role play; collective play; joint play; guided play; guided pretend play; guided dramatic play; play based; purposeful play; playful learning; playful pedagogy. Additionally, the search included terms related to early childhood education contexts: ("preschool" OR "early childhood" OR "kindergarten" OR "early years"). For PubMed, the main search terms were entered in quotation marks to focus on exact matches of the keywords. Furthermore, a "filter by age" was applied to limit results to studies involving young children, specifically those within the age range relevant to preschool settings. To supplement the database searches, a total of 39 sources from a systematic review on guided play (Skene et al., 2022) were included. These sources were specifically selected due to their relevance and focus on guided play strategies in early childhood education settings. A list of all studies identified for inclusion was shared with an expert panel to ensure that key papers had not been missed. Any additional papers suggested were screened for inclusion. Reference list checking (backward citation searching) of all studies identified for inclusion was also conducted. Searches were conducted on August 8-10, 2025.

Data Collection and Analysis

Study Selection and Data Extraction

All records retrieved through the search were exported to Rayyan, an AI-assisted systematic review platform, and screened by titles and abstracts against predefined eligibility criteria. Disagreements were resolved through discussion. Eligible or unclear records were then reviewed at the full-text level. Data extraction focused on participant demographics, study characteristics, intervention details, required teacher skills and support, as well as challenges for implementation, children and teacher outcomes. Verification of inclusion decisions was conducted by two reviewers.

Risk of Bias in Included Studies

Given the heterogeneity of study designs included in this review, risk of bias was assessed using design-specific tools. For randomized controlled trials, risk of bias was evaluated using the Cochrane Risk of Bias tool version 2 (RoB 2; Sterne et al., 2019), for quasi-experimental studies, including one mixed method study, the Risk of Bias in Non-randomized Studies of Interventions tool (ROBINS-I; Sterne et al., 2016) was used, and the Critical Appraisal Skills Programme checklist for qualitative research (CASP, 2018) was used for qualitative studies. The risk of bias assessment was conducted using an AI agent that systematically analyzed each study report (n = 35), applying standardized criteria. The process was carefully validated by one reviewer through manual checks and calibration.

Narrative Synthesis Approach

A narrative synthesis was conducted to summarize and interpret findings across all included studies (Popay et al., 2006). This approach was chosen due to the substantial heterogeneity of study designs and a large proportion of qualitative studies, which precludes meaningful meta-analytic aggregation. The aim of this review is to provide a comprehensive qualitative understanding of the complexities, challenges, and outcomes related to sociodramatic play integration in early childhood education, which is best captured by a narrative approach that highlights both contextual nuances and measurable outcomes.

Study characteristics, participant details, and intervention features were organized within a conceptual framework structured around five identified implementation approaches. Within each approach, child outcomes were categorized into three domains: content, development, and play maturity; teacher involvement was examined as approach-specific, while barriers to implementation were considered as a cross-cutting dimension shaping outcomes across all approaches.

Qualitative findings were synthesized thematically: from each qualitative study we extracted illustrative data (participant quotations, described play episodes, and researcher observations) and grouped them under the review's analytic dimensions (implementation approaches, teacher competencies, barriers, and outcomes). Representative extracts are reported verbatim to preserve contextual meaning, consistent with the aims of narrative synthesis (Popay et al., 2006).

Results

Study Selection

The systematic search results are displayed in a PRISMA flow diagram (*Figure 1*). In total, the search yielded 1,366 potentially eligible studies; 1,271 studies were identified via electronic search and 95 through hand searching, reference list checking, and recommendations from an expert panel. Following eligibility screening, 95 studies were identified for full-text screening. A total of 35 papers (27 studies) met the inclusion criteria and were included in the review. Reasons for exclusion are noted in the PRISMA flow diagram (see *Figure 1*).

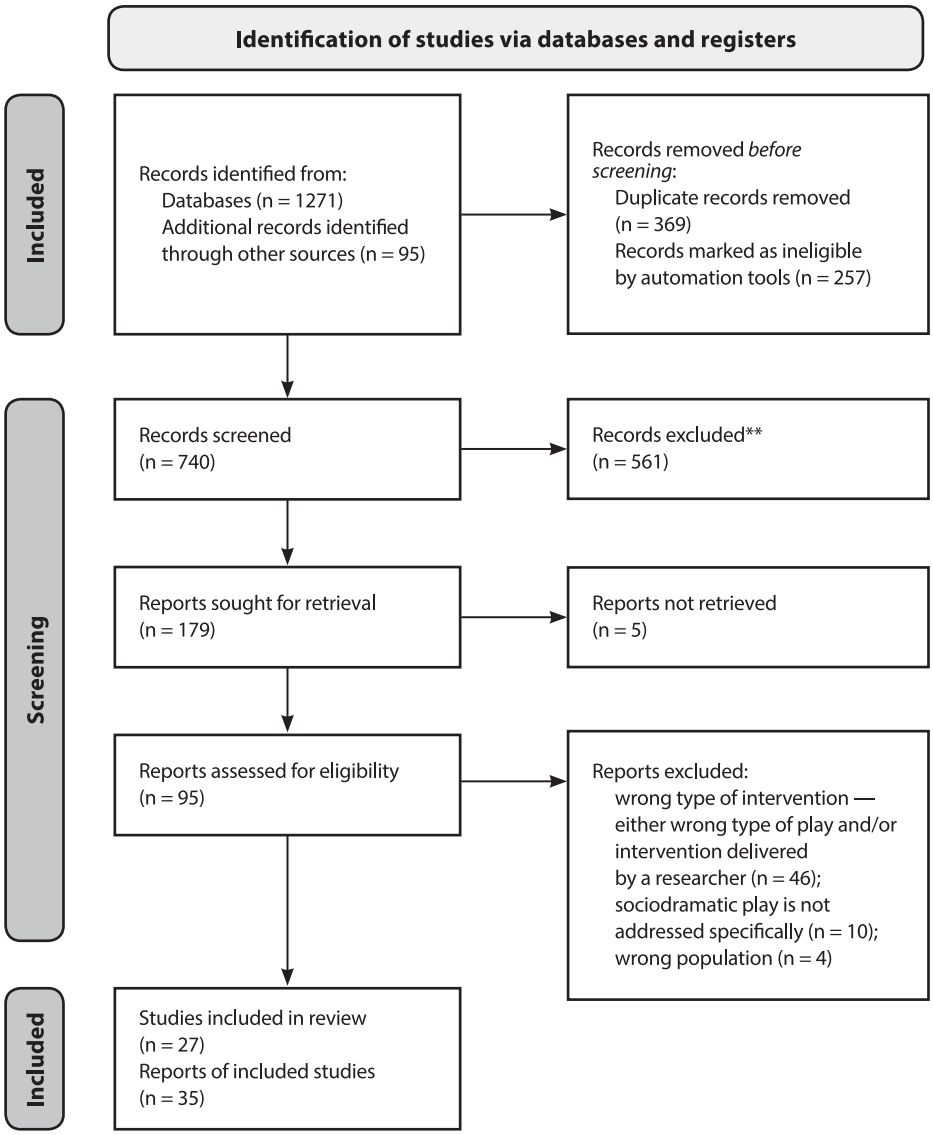

Figure 1. PRISMA flowchart of search results, included and excluded studies

Included Studies Description

Only one study (Dickinson et al., 2018) employed a cluster randomized controlled trial design, with classrooms randomly assigned to conditions. The search also identified 8 quasi-experimental studies and 18 qualitative studies. Of these, 11 qualitative studies described in 18 papers explore Conceptual PlayWorlds developed by M. Fleer and conducted under her guidance or with her participation. See *Table 1* for summary of study characteristics.

Table 1

Included Study Characteristics

Author, year	Study design	Country	Total number of participants (children/ teachers)	Number/ duration/ period of play sessions	Main outcome(s)
Different strategies of sociodramatic play implementation					
Baumer et al., 2005	QE	USA	38/2	14 sessions/ 2 h/ 14 weeks	Narrative comprehension; narrative production.
Brėdikytė, 2022	Q	Lithuania	300+/26	60 play episodes/ 6 months	Changes in behavior, relationship, motivation, attitudes, self-regulation.
Dickinson et al., 2018	Cluster RCT	USA	227/16	12 sessions/ 15–20 min/ 12 weeks	Receptive vocabulary knowledge; expressive vocabulary; self-regulation (cognitive and behavioral); general vocabulary knowledge.
Edwards, Cutter-Mackenzie, 2013	QE	Australia	114/16	3 sessions	Learning of biodiversity concepts.
Ferholt & Rainio, 2016	Q	USA	20/1	14 sessions/ 2h/ 1 year	Inclusion of previously excluded child.
Gmitrova, 2013	QE	Slovakia	368/92	3 days per play theme/3 themes	Play indicators (cognitive).
Hakkarainen, 2008	Q	Finland	31/3	twice a week/ 5 months	Use of cultural tools (signs, symbols, concepts) in interactions; initiation of activities with others using these cultural tools.
Hakkarainen et al., 2013	Q	Finland Lithuania	110/200+	Multiple sessions once a week/ 3-4 h/ 7-year period of research	Joint play.
Hollenstein et al., 2022; Hollenstein & Vogt, 2024	MM Q	Switzerland	15/not specified	every day for 1 to 1.5 h /3.5 months	Problem-solving processes and strategies; role of the teacher in guided pretend play.
Levy et al., 1986	QE	USA	28/ not specified	15 sessions/ 2 months	Language development (Peabody Picture Vocabulary Test).

Author, year	Study design	Country	Total number of participants (children/ teachers)	Number/ duration/ period of play sessions	Main outcome(s)
Levy et al., 1992	QE	USA	9/2	9/ 30 min/ 9 days	Number of words spoken within a 15-minute period; mean length of minimal terminable unit (MLT); number of specific theme-related words spoken within a 15-minute period; number of concept words (color, shape, number, quantity, space, time) spoken within a 15-minute period
McCabe, 2017	QE	Ireland	37/ not specified	10 sessions	Play skills (Smilansky scale).
Nicolopoulou et al., 2015	QE	USA	216/ not specified	regular activity during a year	Narrative skills, vocabulary, emergent literacy (print and word awareness), pretend abilities; social competence (peer cooperation, self-regulation).
Richard et al., 2021	QE	Switzerland	79/10	11 sessions/ 1h/ 11 weeks	Emotion comprehension; prosocial behavior; negative emotion regulation.
Vartiainen et al., 2024	Q	Finland	12/2	4 sessions/ 42–68 min/ not specified	Learning of sustainability concepts.
Vidal Carulla et al., 2021	Q	Sweden	20/ not specified	15 months	Emerging chemistry; executive functions development.
				Conceptual PlayWorlds	
Chen, 2024; Disney & Li, 2022; Li & Disney, 2021	Q	Australia	22/3 (2)	4-6 sessions/ 1.5-2 h	Mathematical problem-solving skills; understanding of mathematical concepts; teachers' dynamic pedagogical positioning.
Fleer, 2017	Q	Australia	26/3	4 weeks	Scientific learning.
Fleer et al., 2017; Fleer et al., 2020	Q	Australia	91/10 (8)	10 sessions/ 10 weeks	Executive functions development.
Fleer, 2020a; Fleer, 2020b	Q	Australia	13/2	12 weeks/12 months	Development of engineering thinking, concept understanding; problem-solving skills; development of Engineering PlayWorld model.

Author, year	Study design	Country	Total number of participants (children/ teachers)	Number/ duration/ period of play sessions	Main outcome(s)
Fleer, 2021	Q	Australia	31/2	regular playworlds during 2 years	Teacher development.
Fleer et al., 2022	Q	Australia	28*	not specified	Integration of science concepts into play-based settings; development of Conceptual PlayWorld model.
Fleer, 2023a Fleer, 2023b	Q	Australia	18/4 (3)	8 sessions/11 weeks	Development of different forms of imagining (affective, embodied, amplified, collective); creation of motivating conditions for science learning; collective problem-solving in imaginary situations.
Li et al., 2022	Q	China	30/1	weekly/ 1.5 h/ 2 years research period	Teacher developed skills in child-centered problem-solving. Improved confidence in teaching STEM concepts. Enhanced ability to balance play and learning goals.
Ma et al., 2022 Ma et al., 2023 Zhang et al.,2025	Q	China	34/2	8 sessions	Active problem-solving. Personally meaningful mathematical learning, successful integration of repeating pattern concept into play activities. Development of teachers' emotional awareness and pedagogical practices in play-based mathematics teaching, successful integration of new pedagogical approach, improved understanding of children's learning processes.
Meng et al., 2022**	Q	China	10/5	10 weeks	Transformation of raw emotions into conscious feeling states, improved fear regulation in play situations.
Utami et al., 2022	Q	Indonesia	38/10	16 h of video observation	Development of mathematical concepts and self-regulation through play, creation of motivating learning conditions by teachers acting as play partners.

*Note. Q – qualitative study design; QE – quasi-experimental study design; RCT — randomized controlled trial; MM — mixed methods; *teachers only; **orphanage.*

The review includes research reports from countries with diverse early childhood education (ECE) systems. Finland (n = 3), Sweden (n = 1) implement comprehensive publicly-funded ECE systems emphasizing play-based learning, outdoor education, and holistic child development (Garvis & Ødegaard, 2022). Lithuania (n = 2) has transitioned toward child-centered education focusing on competence development while maintaining some structured learning elements (European Commission/EACEA/Eurydice, 2019). Ireland (n = 1) implements the universal Early Childhood Care and Education (ECCE) scheme offering free preschool education (Hayes & O'Neill, 2019). Slovakia (n = 1) emphasizes preparation for primary education with a gradual shift toward inclusive practices and play-based methods (Tománková & Vávrová, 2021). Switzerland (n = 3) maintains a decentralized cantonal system with significant regional variations in provision and curriculum (Petitpierre & Gyger, 2020). The USA (n = 6) demonstrates a market-oriented approach with mixed funding sources and substantial state-level policy differences (OECD, 2021), while Australia (n = 13), where most of Conceptual PlayWorlds studies were conducted, implements a national quality framework supporting play-based pedagogical innovations that bridge imaginative play with intentional teaching (Fleer & Hammer, 2021). China (n = 5) and Indonesia (n = 1) represent rapidly evolving ECE systems with growing emphasis on quality standards amid significant rural-urban disparities in educational resources and approaches (Rao et al., 2021; Sun et al., 2022).

Participant Characteristics

The variation in participant number draws attention to limitations in generalizability and applicability of the sociodramatic play implementation in ECE settings. Over a half the studies (n = 18 of 27) had fewer than 50 child participants, 4 studies included approximately 100 child participants, and only 4 included over 200 child participants. One study reported only teacher participants. In total, the studies involved 1,937 child participants; participant numbers varied from nine (Levy et al., 1992) to 300+ (Brėdikytė, 2022). Some qualitative studies focused on a particular child and other children were addressed only as her playmates (Ferholt & Rainio, 2016; Levy et al., 1992; and most of Conceptual PlayWorld studies). Five studies did not specify the number of teacher participants (Hollenstein et al., 2022; Hollenstein & Vogt, 2024; Levy et al., 1986; McCabe, 2017; Nicolopoulou et al., 2015; Vidal Carulla et al., 2021). Hakkarainen et al. (2013) analyzed findings from a longitudinal project and refer to a sample of 200+ teachers. Gmitrova (2013) reported 92 teacher participants. Seven studies included 10-30 teacher participants and the rest, 2-5 teacher participants (see *Table 1*).

Variation in children's age was limited as an inclusion criteria: most studies (n = 22 of 27) exclusively reported on participants aged 3–6 years. One study included children aged 0–5 (Hakkarainen et al., 2013), three studies included children up to 7 years (Baumer et al., 2005; Fleer, 2023a; Fleer, 2023b; Hakkarainen, 2008) and one – up to 11 (Meng et al., 2022) as the study was carried out in an orphanage.

Four studies focused on children from low-income backgrounds (Baumer et al., 2005; Dickinson et al., 2018; McCabe, 2017; Nicolopoulou et al., 2015); three stud-

ies reported a mix of socio-economic and cultural backgrounds (Edwards & Cutter-Mackenzie, 2013; Fleer et al., 2017; Vartiainen et al., 2024), others either did not specify income level (n = 16) or defined it as middle-class (n = 4). The variety of socio-economic backgrounds highlights the universal nature of sociodramatic play as a way to educate and develop children.

Duration of Exposure

Studies employing sociodramatic play strategies demonstrate significant variability in intervention duration, ranging from short-term implementations (*e.g.*, Dickinson et al., 2018: 12 sessions over 12 weeks) to long-term approaches (*e.g.*, Hakkarainen et al., 2013: a 7-year research period). Session frequency varied considerably across studies, with some adopting daily engagement models (*e.g.*, Hollenstein et al., 2022: daily sessions for 3.5 months) while others followed less intensive schedules (*e.g.*, Hakkarainen et al., 2013: once a week). The duration also shows diversity, from brief 15–30-minute episodes (*e.g.*, Levy et al., 1992) to extended 2-hour sessions (*e.g.*, Baumer et al., 2005; Ferholt & Rainio, 2016).

Conceptual PlayWorld interventions typically involve medium-length implementation periods, with studies ranging from 8 weeks (*e.g.*, Ma et al., 2022) to 12 months (*e.g.*, Fleer, 2020a), emphasizing a more sustained approach to play implementation. Weekly play sessions lasted approximately 1.5–2 hours (*e.g.*, Chen et al., 2024; Disney & Li, 2022).

This variability likely reflects both differences in conceptual frameworks and organizational constraints faced by researchers, as none of the reviewed studies provided explicit justification for their chosen implementation parameters.

Fidelity of Implementation

In the majority of studies, interventions were implemented through structured protocols with clearly defined teacher roles, and implementation fidelity was actively monitored through ethnographic observation, audio recording, or research assistant oversight. Deviations from protocol were managed differently: while Dickinson et al. (2018) documented fidelity issues (*e.g.*, 62% adherence during reading), other studies embraced flexibility, reflecting adaptations intrinsic to collaborative educational designs rather than protocol violations (Chen et al., 2024; Fleer et al., 2017 and others).

Professional Development and Support of Sociodramatic Play Implementation

Teacher training for implementing sociodramatic play in classrooms demonstrated significant variation across studies that addressed this aspect in reports, ranging from minimal to extensive professional development programs (see *Table 2*).

Intensive training programs were exemplified by the Narrative Play and Learning project, which provided a six-month program for in-service teachers led by university researchers and experienced practitioners, with monthly meetings and personal mentors from the research team (Brėdikytė, 2022). Similarly comprehensive was a special three-year university extension program focused on "Narrative learning" (Hakkarainen, 2008). Other studies implemented more modest training approaches,

such as half-day professional development sessions introducing pretend play, content knowledge, and gender-sensitive guided play (Hollenstein et al., 2022; Hollenstein & Vogt, 2024), or brief two-hour group training supplemented with detailed manuals (Nicolopoulou et al., 2015). The content of teacher training included theoretical foundations of play pedagogy, practical skills for observation and documentation (Gmitrova, 2013), specific intervention techniques (Richard et al., 2021), content knowledge across domains including mathematics (Disney & Li, 2022), science (Fleer, 2017), biodiversity concepts (Edwards & Cutter-Mackenzie, 2013), and digital competencies (Hollenstein & Vogt, 2024).

Across all training models, a common element was the provision of practical guidance materials, detailed manuals or planning proformas to support independent implementation. Researchers commonly adopted a collaborative approach to teacher training, involving regular meetings and reflections on practice implementation (*e.g.*, Fleer et al., 2020) or at least targeted feedback (Dickinson et al., 2018). Some studies incorporated video observation and analysis into their training programs, teaching educators how to observe, detect, and record children's playing manifestations through practical video sessions (Brédikyté, 2022). A notable trend across studies was the initial intensive coaching decreasing over time to promote teacher autonomy (Dickinson et al., 2018; Disney et al., 2022; Fleer et al., 2020; McCabe et al., 2017).

In implementing sociodramatic play, specific tools were found to be effective. In Brėdikytė (2022), digital video recordings of initial play sessions served as a reflective tool, allowing teachers to self-assess their interactions with children. Watching these videos highlighted areas of frustration and prompted teachers to adjust their pedagogical approaches, making the recordings a critical cornerstone for professional reflection. Dickinson et al. (2018) introduced guidance cards detailing how to discuss, use key vocabulary, provide definitions, and elicit children's responses. Teachers reported high fidelity in applying these strategies and valued the structured supports, which facilitated more consistent and effective engagement during play sessions.

Risk of Bias in Included Studies

The only RCT (Dickinson et al., 2018) was rated as raising some concerns overall, primarily due to inconsistent intervention delivery during play sessions and differential attrition. The eight quasi-experimental studies varied in overall risk: one was rated critical risk (Levy et al., 1986) due to the complete absence of a control group, three — serious risk (Baumer et al., 2005; Levy et al., 1992; McCabe, 2017), and three — moderate risk (Edwards & Cutter-Mackenzie, 2013; Gmitrova, 2013; Nicolopoulou et al., 2015); one study provided insufficient information for an overall judgment (Richard et al., 2021). Confounding was the predominant source of bias across non-randomized studies. Measurement of outcomes was the strongest domain, with low risk ratings. Missing data bias was typically low, owing to comprehensive video recordings, interviews, and structured data collection, although attrition affected some studies, as in Baumer et al. (2005) with 33% missing data in the control group.

Seven qualitative studies of different strategies of sociodramatic play implementation showed overall methodological quality moderate to high. All studies demonstrated clear and rich presentation of findings, and most provided sufficient justification for their qualitative approach. However, several systematic weaknesses were identified across the corpus. Reflexivity was the most consistent limitation: five of seven studies failed to adequately address the researcher–participant relationship, which is particularly concerning given that researchers frequently occupied dual roles as both intervention designers and data collectors. Analytical rigor was insufficient in five studies, with most lacking interrater reliability procedures or transparent coding protocols. Sampling strategies raised concerns in six studies, with volunteer and convenience samples predominating. Brėdikytė (2022) noted that volunteer teachers often possessed specific characteristics (*e.g.*, higher motivation, openness to innovation) that may not represent the broader teaching population. The 11 qualitative Conceptual PlayWorld studies, reported across 18 papers, were assessed using a cluster appraisal approach rather than individual ratings, as the studies share a common theoretical framework and methodological design. Strengths of the cluster included rich multi-source data collection and clear, theoretically grounded presentation of findings. However, the same weaknesses were identified regarding reflexivity, analytical rigor, and sampling strategies.

Narrative Synthesis

Sociodramatic Play Implementation Approaches

Five approaches were identified, each differing in the degree and features of teacher involvement, integration of curriculum content, and structure of narrative and play (see *Table 2*).

Storytelling and story-acting (STSA) is a structured, child-centered curriculum module designed to promote preschoolers' narrative, literacy, and social skills. It is a particular pedagogical model which refers specifically to Paley's (1990) structured dictation-and-enactment module. During "choice time," children voluntarily dictate stories to a teacher, who writes the stories verbatim with minimal intervention, occasionally asking clarifying questions. Each child then selects a role for themselves and chooses peers to act out the story. The teacher reads the story aloud and supervises its dramatization, ensuring all dictated stories are enacted in group time, fostering narrative sharing and collaborative engagement. The approach is grounded in Vygotskian theory, and integrates storytelling and pretend play to support both cognitive and socio-emotional development. For example, in Nicolopoulou et al. (2015), a teacher wrote down each child's self-initiated, verbatim story with minimal intervention, occasionally asking clarifying questions such as "What happened next?" or "Is that the end?" During group time, the teacher read every story aloud while the child-author, having chosen a role and selected peers for the remaining parts, enacted it before the assembled class. Each child rotated through three roles: story-composer, actor, and audience member.

Playworlds are the immersive, co-created narrative worlds where children enact roles and solve problems within a structured story. Teachers facilitate entry and exit rituals, maintain narrative coherence, scaffold role-taking, and introduce problem-

Table 2

Key Features of Sociodramatic Play Approaches

Approach	Outcome domain	Teacher involvement	Teacher support strategies
Storytelling and story-acting (STSA)	**Development-focused** (Nicolopoulou et al., 2015)	Record children's stories, supervise enactment, minimal prompting	Brief PD sessions; short orientation or guidance materials
Playworlds	**Development-focused** (Baumer et al., 2005; Brėdikytė, 2022; Hakkarainen, 2008; Hakkarainen et al., 2013; Ferholt, Rainio, 2016) **Maturity of play as an outcome** (Hakkarainen, 2008; Hakkarainen et al., 2013; Ferholt, Rainio, 2016)	Facilitate entry/exit, scaffold roles, maintain narrative, introduce problem-solving	Intensive multi-week program; mentoring and structured meetings
Conceptual PlayWorlds (CPW)	**Content-focused** (Chen, 2024; Disney & Li, 2022; Fleer, 2017; Fleer, 2020a; Fleer, 2020b; Fleer et al., 2022; Fleer, 2023a; Fleer, 2023b; Li et al., 2022 Li & Disney, 2021; Ma et al., 2022; Ma et al., 2023; Utami et al., 2022; Zhang et al.,2025) **Development-focused** (Fleer et al., 2017; Fleer et al., 2020; Meng et al., 2022*; Utami et al., 2022)	Co-create storyworld, pose challenges, guide learning, scaffold collaboration	Long, structured training (weeks/months); stepwise guidance and iterative reflection
Play scaffolding (*e.g.*, using drama or play materials)	**Content-focused** (Edwards, Cutter-Mackenzie, 2013) **Development-focused** (Hollenstein et al., 2022; Hollenstein & Vogt, 2024 Levy et al., 1986; Levy et al., 1992) **Maturity of play as an outcome** (Gmitrova, 2013; McCabe, 2017)	Provide props, model roles, guide scenarios, gradually release control	Moderate PD sessions; manuals, stepwise coaching, feedback sessions
Playful components in learning (play-based learning)	**Content-focused** (Dickinson et al., 2018 Richard et al., 2021 Vartiainen et al., 2024 Vidal Carulla et al., 2021)	Set playful context, link to academic tasks, rotate roles, support reflection	Short PD sessions or workshops; manuals, structured activity cards, moderate guidance

Note. PD – professional development

solving challenges linked to the story. Props, cultural tools, and negotiated storylines are employed to sustain engagement and deepen learning. For example, in Ferholt and Rainio's (2016) Narnia Playworld, the teacher entered in a role as the White Witch and wove a previously excluded child's disruptive proposal (to "blow up the

whole school") into the collective narrative, so that participation in the shared story became the vehicle for the child's inclusion.

Conceptual PlayWorlds (CPW) integrate curriculum concepts such as mathematics, science, or engineering into narrative play. Teachers and children collaboratively create an immersive narrative environment, drawing on storybooks, children's interests, or invented characters, and dramatize the plot together. The approach involves five key steps: selecting a story to engage children emotionally; designing a PlayWorld space; using structured routines for entering and exiting the PlayWorld; planning play inquiries or problem scenarios that challenge children's skills and knowledge; and planning teacher interactions to scaffold conceptual learning in role. Educators take active roles as co-players, guiding the narrative, posing challenges, and ensuring opportunities for collaboration, engagement, and problem-solving. For example, in Zhang et al. (2025), teachers and children adapted a chapter of Journey to the West, entering the imaginary world through "a waterfall cave made of colored paper" and dramatizing a sequence of conceptual problems in role: rescuing the fake Monkey King (direction and orientation), crafting a weapon of matching height (measurement), and decorating it so that no three strips of one color appeared together (repeating patterns).

Play scaffolding (*e.g.*, using drama or play materials) emphasizes teacher facilitation of pretend play with props, dramatization, and structured scenarios. Teachers select themes, provide materials and visual cues, model role-play interactions, and gradually delegate responsibility to the children while observing and supporting collaborative enactment. Special strategies include stepwise guidance, role rotation, and reflection discussions to consolidate cognitive and social skills. For example, in play-scaffolding studies, teachers selected a theme, provided props and visual cues, modelled role-play interactions, and gradually delegated responsibility to the children (*e.g.*, Gmitrova, 2013; McCabe, 2017).

Playful components in learning (play-based learning) are used as a context for academic tasks. Teachers construct playful learning environments, using props, toys, soundscapes, and scenario-based narratives to engage children in problem-solving and co-creation activities. Rotating roles, structured reflection, and linking playful activities to content objectives are emphasized. For example, in Dickinson et al. (2018), pretend play served as a structured context for vocabulary instruction: teachers used guidance cards specifying target words, definitions, and prompts to elicit children's responses across a fixed sequence of short, scripted sessions. Play here functions as a vehicle for pre-defined academic content rather than as an open-ended imaginary world.

The results indicate that some approaches, such as Playworlds and Conceptual PlayWorlds, require extensive preparation and high levels of teacher qualification, which is why teachers typically participate on a voluntary basis, with significant support from researchers. While the results of these approaches are valuable, they are often qualitatively documented and cannot be objectively measured. On the other hand, highly structured approaches with clearly measured outcomes implementing playful components in learning tend to struggle with fidelity issues, which is also problematic.

Pedagogical Competencies and Skills

Implementing sociodramatic play in preschool requires educators to master a range of pedagogical skills and competencies. Teachers must be able to scaffold children's play, extend role speech, monitor behaviors, and intervene sensitively without dominating the play (McCabe, 2017; Nicolopoulou et al., 2015; Richard et al., 2021). Expertise in narrative facilitation is essential, including co-constructing stories, improvising within play scenarios, and assuming dramatic roles to maintain engagement (Brėdikytė, 2022; Hakkarainen, 2008, 2013; McCabe, 2017). Educators require domain-specific knowledge, such as STEM, science, or digital literacy, to embed curriculum concepts meaningfully into play (Fleer, 2020a, 2020b; Hollenstein & Vogt, 2024). Competency in dynamic pedagogical positioning allows teachers to shift between observation, guidance, and co-play according to the children's developmental and motivational needs (Disney & Li, 2022; Li et al., 2021; Utami et al., 2023). This is illustrated in Disney and Li's (2022) analysis of a "Mathematical Playworld" based on the story "Room on the Broom." Playing a witch and a dog, the teachers first took a "below" position, observing as the children built a broom for 13 characters but focused only on their own seats, "displaying a distinct lack of mathematical thinking or language" (p. 8). They then shifted to an "equal" position: "I've noticed that there are bigger seats and small seats," one teacher remarked, while another asked, "So, does everyone have the same amount of room?" (p. 6). When one child still equated each block with one seat regardless of length, the teachers moved to an "above" position, physically seating children on long and short blocks to make the measurement concept concrete. This dynamic shifting enabled intentional concept teaching without abandoning the imaginary situation.

Social-emotional skills, including listening, turn-taking, and coordinating group actions, are critical for sustaining shared imaginative situation and motivating engagement (Brėdikytė, 2022; Meng et al., 2022; Utami et al., 2023). Teachers also need skills to facilitate higher-order thinking and executive function development within play (Fleer, 2020b; Fleer et al., 2020; Fleer, 2023b). Reflective practice and adaptability are essential, as educators must adjust strategies to children's interests and developmental progress, co-constructing meaningful play experiences while integrating learning goals (Fleer, 2021; Ma et al., 2022, 2023; Nicolopoulou et al., 2015). Across the reviewed studies, effective implementation of sociodramatic play emerged as a multidimensional professional competency, encompassing narrative and dramatic skills, dynamic pedagogical positioning, domain-specific knowledge, social-emotional facilitation, and reflective adaptability.

Challenges and Limitations for Sociodramatic Play Implementation

Teachers encounter multiple barriers — personal, pedagogical, and systemic – that impede the effective implementation of sociodramatic play. At the personal level, educators frequently lack improvisation skills essential for spontaneous and authentic play participation (Brėdikytė, 2022). Many experience anxiety in unpredictable situations, demonstrate a persistent need to control activities, lack patience, rush sessions, and inadvertently overshadow children's initiatives (Brėdikytė, 2022). Transitioning from observer to active participant in children's play is particularly

difficult, with some teachers perceiving engagement in play as contrary to their professional identity (Dickinson et al., 2018). Teachers' own words captured this shift vividly. Reviewing video of their first play sessions, one teacher admitted: "I didn't notice it [my behavior] until I started filming […] when I watched the footage, I realized, it's wrong!" (Brėdikytė, 2022, p. 599). Teachers listed the behaviors they struggled to overcome: "a constant desire to control and guide children," "lack of patience," "continuous rush," "seeking the result," "overshadowing the children" (Brėdikytė, 2022, p. 599) and identified "leaving the comfort zone" by "taking a role and playing with children" as the hardest step (p. 599). In some contexts, such as in China, cultural tendencies to conceal personal feelings further limit the open dialogue and self-reflection necessary for adopting new pedagogical approaches (Zhang et al., 2025). The emotional weight of pedagogical change is illustrated by Teacher Qing, who, after a session in which she could not respond to children's unexpected difficulties in measurement, shed tears during the collective reflection: "I feel like I am dragging everyone down, even with all the preparation we have done, but I still need to improve my ability to respond to my children spontaneously and not just teach them mathematics concepts. At that moment … I did not know what to do as a character" (Zhang et al., 2025, p. 9). She traced this difficulty to her habitual practice: "I normally follow a structured lesson plan to teach mathematics.… One disadvantage of placing extensive reliance on teachers' planned instructions is that it fails to acknowledge children's abilities and their perspectives" (Zhang et al., 2025, p. 9).

Pedagogically, many educators report feeling inadequately prepared and lacking confidence, particularly regarding digital competencies needed to guide children's understanding of technological concepts (Hollenstein & Vogt, 2024). Similarly, teachers often express limited confidence in their scientific knowledge (Fleer, 2023) and frequently avoid teaching science due to insufficient expertise (Fleer et al., 2022). Engineering concepts pose additional challenges, as teachers often struggle to recognize engineering affordances within play (Fleer, 2020). Beyond content knowledge, teachers face the challenge of balancing structured guidance with child-led exploration. Excessive or inappropriate engagement can disrupt the flow of children's play and agency (Hollenstein, 2024). This tension is especially pronounced in mathematical learning, where educators must introduce concepts without reverting to traditional instructional methods (Disney & Li, 2022). Teachers often struggle to determine when to step aside or fully participate (Vartiainen, 2024) and frequently misunderstand pretend play as merely a resting activity or as beneficial only when it emerges spontaneously from children rather than through adult facilitation (Gmitrova, 2013; McCabe, 2017).

Systemic challenges further complicate implementation. Time constraints limit opportunities for teachers to plan play sessions and to work with multiple small groups (Dickinson et al., 2018), and competing academic demands create tension between play-based approaches and traditional curriculum expectations (Baumer et al., 2005; McCabe, 2017). Implementation fidelity varies, with some educators embracing new approaches enthusiastically while others demonstrate resistance or inconsistency (Nicolopoulou et al., 2015).

Child Development Outcomes

The outcomes reported across the included studies were categorized into three domains: content-focused, development-focused, and maturity of play as an outcome (see *Table 2*). Content-focused outcomes refer to the acquisition of knowledge or skills in specific areas, such as vocabulary, mathematical concepts, scientific understanding, or sustainability learning. For instance, Dickinson et al. (2018) reported gains in receptive and expressive vocabulary after 12 sessions of 15–20 minutes each over 12 weeks, while Edwards and Cutter-Mackenzie (2013) measured learning of biodiversity concepts across three sessions. Development-focused outcomes targeted children's overall developmental progress, encompassing self-regulation, executive functions, and social competence. Nicolopoulou et al. (2015) examined social competence and self-regulation, while Fleer et al. (2017) focused on executive functions development. Richard et al. (2021) specifically measured emotion comprehension and prosocial behavior. The maturity of play as an outcome was examined through indicators of play development, sophistication, collaboration, or creative engagement during play (Ferholt, Rainio, 2016; Gmitrova, 2013; Hakkarainen, 2008; Hakkarainen et al., 2013; McCabe, 2017). In a multilingual group, one teacher observed a "'burst' of children's language … and sociodramatic role-play appeared with lots of interactions among children from different linguistic backgrounds" (Brėdikytė, 2022, p. 602).

Positive effects were consistently reported across multiple studies and all the outcome domains. However, some results were inconsistent or limited in effect. Baumer (2005) found no statistically significant improvement in linguistic complexity following Playworld interventions, even though other language measures showed positive gains. The study by Vidal Carulla et al. (2021) acknowledged limitations in their research design for exploring working memory development, noting that while attention shifting and inhibitory control showed clear benefits, working memory outcomes were less conclusive. Dickinson et al. (2018) found that while both play-based and read-only conditions showed significant vocabulary gains, the addition of play did not produce significantly better outcomes than reading alone. Some Conceptual PlayWorld studies highlighted challenges in consistently achieving pedagogical goals, with teacher experience and engagement influencing outcomes across sessions (Fleer et al., 2022; Ma et al., 2023).

Teacher Development Outcomes

Across the studies emphasizing teacher outcomes, changes in pedagogical practices, professional skills, and emotional engagement were documented. Teachers demonstrated improvements in scaffolding strategies, role adoption, and facilitation of child-centered problem solving (Li et al., 2022; Utami et al., 2023; Zhang et al., 2025). Increased confidence, emotional awareness, and greater adaptability in play facilitation were also reported (Fleer, 2021; Fleer et al., 2022; Ma et al., 2022, 2023). Longitudinal studies highlighted improvements in the integration of curriculum content into play and in creating motivating learning environments (Brėdikytė, 2022; Fleer, 2020b; Hakkarainen et al., 2013).

Discussion

Summary of Evidence

This systematic review synthesized evidence from 27 studies (described in 35 papers) examining sociodramatic play strategies in ECE settings across diverse international contexts. The review provides a clear classification of sociodramatic play implementation approaches, identifying key teacher competencies needed for successful implementation, and highlighting barriers faced by educators. These findings directly address the identified gaps, such as the lack of systematic implementation of play in educational practices.

The review revealed five main implementation models: Storytelling and Story-Acting (STSA), Playworlds, Conceptual PlayWorlds, Play Scaffolding, and Playful Components in Learning. These approaches are described not only in experimental studies but also in theoretical articles (Novlyanskaya, 2025; Veresov et al., 2025). Child development outcomes were broadly positive across three domains: content, development, and maturity of play itself. However, some null findings (Baumer, 2005; Dickinson et al., 2018; Fleer et al., 2022; Ma et al., 2023; Vidal Carulla et al., 2021) are methodologically important: they suggest that the benefits of sociodramatic play may be domain-specific and that the mere presence of play does not guarantee learning and development.

Teacher outcomes included enhanced pedagogical practices, improved professional skills, and deeper emotional engagement with children's play and learning processes.

The review identified essential pedagogical competencies required for effective implementation that strongly resonate with Vygotskian notions of the teacher as a more knowledgeable other who operates within the child's zone of proximal development, providing support that is responsive to moment-by-moment shifts in children's engagement and understanding (Vygotsky, 2016). Teachers must be able to enter imaginary situations authentically, sustain narrative coherence while reaching educational goals, and model the metacommunicative negotiations that characterize mature pretend play. This performative competence is distinct from traditional instructional skills and may explain why many teachers report low confidence in implementing sociodramatic play despite general pedagogical expertise. This tension is not merely technical but reflects deeper theoretical questions about the nature of play and learning (Serafimovich & Zakharova, 2025). Significant barriers to implementation were noted, spanning personal limitations, pedagogical challenges, and systemic constraints.

The present review's findings converge with Skene et al.'s (2022) systematic review and meta-analysis of guided play in educational contexts, which found that guidance during play can enhance children's learning and development. The present review extends this work by specifying the forms that guidance takes in sociodramatic play contexts and by highlighting the importance of teacher competencies.

Implications for Research and Practice

There is a clear need for more rigorous research designs, including larger sample sizes, randomized controlled trials with appropriate controls, and consistent report-

ing of intervention parameters to strengthen evidence quality. Researchers should develop and validate specific instruments for measuring the quality of sociodramatic play implementation and its distinct contributions to child development, beyond general measures of play quality. Extended studies tracking both implementation fidelity and developmental outcomes over multiple years would significantly enhance understanding of long-term impacts and sustainability. Research examining how sociodramatic play approaches can be effectively adapted across diverse cultural and educational contexts is needed, particularly investigating how cultural values influence implementation processes.

Effective implementation of sociodramatic play requires investment in teacher professional development. More detailed investigation into how teachers develop competence in sociodramatic play facilitation over time would inform more effective professional development approaches. Education programs should incorporate specific training in narrative facilitation, improvisation skills, and disciplinary knowledge integration through play, addressing the identified competency gaps. Educational systems should develop comprehensive implementation frameworks providing educators with structured guidance, observation tools, and reflection protocols to support consistent application. Establishing collaborative networks of practitioners implementing sociodramatic play could address the confidence and competency barriers identified, facilitating peer learning and ongoing support. Fostering a culture that values reflective practice and open dialogue is critical to overcoming personal and systemic barriers, enabling teachers to implement meaningful, developmentally appropriate sociodramatic play.

Conclusion

Despite the substantial body of evidence supporting the use of play for developmental and educational purposes, there is a notable gap in research demonstrating how teachers working across various curricula can effectively implement collective sociodramatic play in their daily practice. This disconnect between theoretical understanding and practical application represents a significant challenge for educational settings. Key findings highlight that widespread implementation of sociodramatic play requires approaches that are both practically accessible and intrinsically motivating to teachers, while reliably generating observable and impactful results for children.

This review synthesized evidence from diverse methodological approaches, educational systems, and cultural contexts, providing a nuanced understanding of play implementation globally. Detailed examination of pedagogical approaches, required teacher competencies, and implementation challenges offers practical insights for practitioners. The proposed typology of five implementation approaches provides conceptual clarity in a field characterized by terminological inconsistency, facilitating more precise cross-study comparison. By incorporating both quantitative and qualitative studies, the review captures both the measurable impacts and the rich contextual factors influencing implementation effectiveness. Unlike many educational reviews that focus solely on child outcomes, this study explicitly addressed teacher development as a critical component of successful implementation.

Future research in this area should focus on several directions. Researchers need to explore how sociodramatic play can be effectively adapted across diverse cultural contexts, paying special attention to the influence of cultural values and educational systems on implementation processes and outcomes. It is also important to investigate the most effective models of professional development for teachers, including optimal training duration, content, and follow-up support strategies that would help build competence in sociodramatic play facilitation. Finally, there is a need to develop specific assessment tools to measure the quality of sociodramatic play implementation and its unique contribution to child development, going beyond general play measures.

Limitations

Limitations include high risk of bias due to the non-randomized designs, small and heterogeneous samples, and selection bias in teacher recruitment, potentially limiting the generalizability of findings. The studies varied considerably in how they reported intervention duration, frequency, and intensity, making direct comparisons challenging. An analysis of potential moderators, such as teacher experience, cultural context, and child age, was not conducted in this review due to the heterogeneity of the included studies and the lack of consistent reporting on these variables across the studies. A significant proportion of qualitative studies (11 of 18) focused on Conceptual PlayWorlds developed by M. Fleer, potentially overrepresenting this specific approach.

A further limitation concerns the language and database scope of the search. Restricting eligibility to English-language articles indexed in ScienceDirect and PubMed excluded research published in Spanish and Portuguese and in regional databases (*e.g.*, SciELO, Redalyc, LILACS). Consequently, the substantial body of Latin American scholarship on the development of children's play activity is not represented in this synthesis, which limits the geographical and cultural generalizability of the findings.

Author Contributions

V.S. and N.V. conceived of the idea. V.S. designed the systematic review protocol and performed the literature search. V.S. and M.G. screened the studies and extracted the data. N.V. verified the study selection criteria and the data extraction process, and provided methodological oversight for the systematic review approach. V.S. and M.G. conducted the synthesis of findings and interpreted the results in the context of early childhood education. N.V. contributed critical insights to the theoretical framing of teacher implementation of sociodramatic play. All authors discussed the results and contributed to the final manuscript.

Conflict of Interest

The authors declare no conflict of interest.

References

Baumer, S., Ferholt, B., & Lecusay, R. (2005). Promoting narrative competence through adult–child joint pretense: Lessons from the Scandinavian educational practice of playworld. *Cognitive Development, 20*(4), 576–590. https://doi.org/10.1016/j.cogdev.2005.08.003

Bodrova, E., & Leong, D. J. (2007). *Tools of the mind: The Vygotskian approach to early childhood education* (2nd ed.). Pearson.

Bodrova, E.V., & Leong, D.J. (2025). Ostrovok amplifikatsii v okeane akseleratsii [An island of amplification in the ocean of acceleration]. *Natsionalnii Psikhologicheskii Zhurnal* [National Psychological Journal], *20*(3), 29–41. https://doi.org/10.11621/npj.2025.0303

Bower, C. A., Zimmermann, L., Verdine, B. N., Spiewak Toub, T., Hirsh-Pasek, K., & Michnick Golinkoff, R. (2025). What does play have to do with it? A concrete and digital spatial intervention with 3-year-olds predicts spatial and math learning. *Developmental Psychology, 61*(3), 461–481. https://doi.org/10.1037/dev0001904

Bredikyte, M. (Ed.). (2017). *Narrative environments for play and learning (NEPL): Guidelines for kindergarten and school teachers working with 3–8 year-old children.* Publishing House of the Lithuanian University of Educational Sciences.

Brėdikytė, M. (2022). Adult participation in the creation of narrative playworlds: Challenges and contradictions. *International Journal of Early Years Education, 30*(3), 595–609. https://doi.org/10.1080/09669760.2022.2107895

Bulunuz, M. (2013). Teaching science through play in kindergarten: Does integrated play and science instruction build understanding? *European Early Childhood Education Research Journal, 21*(2), 226–249. https://doi.org/10.1080/1350293X.2013.789195

Chen, H., Disney, L., & Li, L. (2024). Studying children's motives in mathematical problem-solving during transition from kindergarten to school: A Conceptual PlayWorld approach. *Learning, Culture and Social Interaction, 48,* 100850. https://doi.org/10.1016/j.lcsi.2024.100850

Critical Appraisal Skills Programme. (2018). CASP qualitative studies checklist. https://casp-uk.net/casp-tools-checklists/

DePascale, M., & Ramani, G. B. (2024). Promoting children's mathematical and statistical understanding through parent-child math games. *Cognitive Development*, *71*, 101480. https://doi.org/10.1016/j.cogdev.2024.101480

Diamond, A., Barnett, W.S., Thomas, J., & Munro, S. (2007). Preschool program improves cognitive control. *Science, 318*(5855), 1387–1388. https://doi.org/10.1126/science.1151148

Dickinson, D.K., Collins, M.F., Nesbitt, K., Toub, T.S., Hassinger-Das, B., Hadley, E.B., Hirsh-Pasek, K., & Golinkoff, R.M. (2018). Effects of teacher-delivered book reading and play on vocabulary learning and self-regulation among low-income preschool children. *Journal of Cognition and Development, 20*(2), 136–164. https://doi.org/10.1080/15248372.2018.1483373

Disney, L., & Li, L. (2022). Above, below, or equal? Exploring teachers' pedagogical positioning in a playworld context to teach mathematical concepts to preschool children. *Teaching and Teacher Education, 114,* 103706. https://doi.org/10.1016/j.tate.2022.103706

Edwards, S., & Cutter-Mackenzie, A. (2013). Pedagogical play types: What do they suggest for learning about sustainability in early childhood education? *International Journal of Early Childhood, 45*(3), 327–346. https://doi.org/10.1007/s13158-013-0082-5

European Commission/EACEA/Eurydice. (2019). *Key data on early childhood education and care in Europe: 2019 edition* (Eurydice Report). Publications Office of the European Union. https://doi.org/10.2797/966808

Ferholt, B., & Rainio, A.P. (2016). Teacher support of student engagement in early childhood: Embracing ambivalence through playworlds. *Early Years, 36*(2), 114–128. https://doi.org/10.1080/09575146.2016.1141395

Fleer, M. (2017). Scientific playworlds: A model of teaching science in play-based settings. *Research in Science Education, 49*(5), 1257–1278.

Fleer, M. (2020a). Engineering playworld: A model of practice to support children to collectively design, imagine and think using engineering concepts. *Research in Science Education, 52*(2), 583–598. https://doi.org/10.1007/s11165-020-09970-6

Fleer, M. (2020b). Studying the relations between motives and motivation: How young children develop a motive orientation for collective engineering play. *Learning, Culture and Social Interaction, 24,* 100355. https://doi.org/10.1016/j.lcsi.2019.100355

Fleer, M. (2021). An educational experiment into how to bring discipline concepts into play: How a theoretical problem acts as a source of teacher development. *Moscow University Psychology Bulletin. Series 14, 14*(4), 72–103. https://doi.org/10.11621/vsp.2021.04.03

Fleer, M. (2023a). The role of imagination in science education in the early years under the conditions of a Conceptual PlayWorld. *Learning, Culture and Social Interaction, 42,* 100753. https://doi.org/10.1016/j.lcsi.2023.100753

Fleer, M. (2023b). Are we play-proofing preschools, kindergartens, and schools? Conceptual PlayWorlds as the source of children's development of imagination and learning. *American Journal of Play, 15*(3), 302–332.

Fleer, M., Fragkiadaki, G., & Rai, P. (2022). The place of theoretical thinking in professional development: Bringing science concepts into play practice. *Learning, Culture and Social Interaction, 32,* 100591. https://doi.org/10.1016/j.lcsi.2021.100591

Fleer, M., & Hammer, M. (2021). Conceptual playworlds: Foregrounding imagination and creativity as foundational for children's learning. In L. Arnott & K. Wall (Eds.), *Research through play: Participatory methods in early childhood* (pp. 45–61). SAGE Publications.

Fleer, M., Veresov, N., Harrison, L., & Walker, S. (2017). Working with teachers' pedagogical strengths: The design of executive function activities for play-based programs. *Australasian Journal of Early Childhood, 42*(4), 47–55. https://doi.org/10.23965/AJEC.42.4.06

Fleer, M., Walker, S., White, A., Veresov, N., & Duhn, I. (2020). Playworlds as an evidenced-based model of practice for the intentional teaching of executive functions. *Early Years*, *42*(4–5), 572–586. https://doi.org/10.1080/09575146.2020.1835830

Garvis, S., & Ødegaard, E.E. (Eds.). (2022). *Nordic perspectives on early childhood education and care.* Routledge. https://doi.org/10.4324/9781003159124

Gmitrova, V. (2013). Teaching to play performing a main role: Effective method of pretend play facilitation in preschool-age children. *Early Child Development and Care, 183*(11), 1705–1719. https://doi.org/10.1080/03004430.2012.746970

Hakkarainen, P. (2008). The challenges and possibilities of a narrative learning approach in the Finnish early childhood education system. *International Journal of Educational Research, 47*(5), 292–300. https://doi.org/10.1016/j.ijer.2008.12.008

Hakkarainen, P., Brėdikytė, M., Jakkula, K., & Munter, H. (2013). Adult play guidance and children's play development in a narrative play-world. *European Early Childhood Education Research Journal, 21*(2), 213–225. https://doi.org/10.1080/1350293X.2013.789189

Hayes, N., & O'Neill, S. (2019). Little changes, big results: The impact of simple changes to early years learning environments. *Early Years, 39*(1), 64–79. https://doi.org/10.1080/09575146.2017.1342223

Hollenstein, L., Thurnheer, S., & Vogt, F. (2022). Problem solving and digital transformation: Acquiring skills through pretend play in kindergarten. *Education Sciences, 12*(2), 92. https://doi.org/10.3390/educsci12020092

Hollenstein, L., & Vogt, F. (2024). Digital education through guided pretend play. *Learning and Instruction, 93*, 101945. https://doi.org/10.1016/j.learninstruc.2024.101945

Levy, A.K., Schaefer, L., & Phelps, P.C. (1986). Increasing preschool effectiveness: Enhancing the language abilities of 3- and 4-year-old children through planned sociodramatic play. *Early Childhood Research Quarterly, 1*(2), 133–140. https://doi.org/10.1016/0885-2006(86)90024-4

Levy, A.K., Wolfgang, C.H., & Koorland, M.A. (1992). Sociodramatic play as a method for enhancing the language performance of kindergarten-age students. *Early Childhood Research Quarterly, 7*(2), 245–262. https://doi.org/10.1016/0885-2006(92)90007-L

Li, L., & Disney, L. (2021). Young children's mathematical problem solving and thinking in a play-world. *Mathematics Education Research Journal, 35*(1), 23–44. https://doi.org/10.1007/s13394-021-00373-y

Li, L., Fleer, M., & Yang, N. (2022). Studying teacher professional development: How a Chinese kindergarten teacher brings play practices into the program. *Early Years, 42*(1), 104–118. https://doi.org/10.1080/09575146.2021.2000942

Lillard, A.S., Lerner, M.D., Hopkins, E.J., Dore, R.A., Smith, E.D., & Palmquist, C.M. (2013). The impact of pretend play on children's development: A review of the evidence. *Psychological Bulletin, 139*(1), 1–34. https://doi.org/10.1037/a0029321

Ma, Y., Fleer, M., Li, L., & Wang, Y. (2023). A cultural historical study of how a conceptual playworld creates conditions for personally meaningful mathematics in the Chinese kindergarten context. *European Early Childhood Education Research Journal, 31*(6), 850–865. https://doi.org/10.1080/1350293X.2023.2200020

Ma, Y., Wang, Y., Fleer, M., & Li, L. (2022). Promoting Chinese children's agency in science learning: Conceptual PlayWorld as a new play practice. *Learning, Culture and Social Interaction, 33,* 100614. https://doi.org/10.1080/1350293X.2023.2200020

Marulis, L.M., & Neuman, S.B. (2010). The effects of vocabulary intervention on young children's word learning: A meta-analysis. *Review of Educational Research, 80*(3), 300–335. https://doi.org/10.3102/0034654310377087

McCabe, U. (2017). The drama in sociodramatic play: Implications for curriculum and pedagogy. *NJ: Drama Australia Journal, 41*(1), 3-13. https://doi.org/10.1080/14452294.2017.1329689

Meng, X., Fleer, M., Li, L., & Hammer, M. (2022). A cultural-historical study of the development of emotion regulation through collective play in an institutional care setting in China. *Learning, Culture and Social Interaction, 32,* 100590. https://doi.org/10.1016/j.lcsi.2021.100590

Smits-van der Nat, M., Wilt, F. van der Meeter, M., & van der Veen, C. (2024). The value of pretend play for social competence in early childhood: A meta-analysis. *Educational Psychology Review, 36*(2), 46. https://doi.org/10.1007/s10648-024-09884-z

Nicolopoulou, A., Cortina, K.S., Ilgaz, H., Cates, C.B., & de Sá, A.B. (2015). Using a narrative- and play-based activity to promote low-income preschoolers' oral language, emergent literacy, and social competence. *Early Childhood Research Quarterly, 31*, 147–162. https://doi.org/10.1016/j.ecresq.2015.01.006

Novlyanskaya, Z.N. (2025). Perekhod ot igry k ucheniiu v kurse literaturnogo chteniia v sisteme razvivaiushchego obucheniia D.B. El'konina i V.V. Davydova. [Transition from play to learning in the course of literary reading in the system of developmental education by D.B. Elkonin and V.V. Davydov]. *Vestnik Moskovskogo Universiteta. Seriia 14. Psikhologiia* [Moscow University Psychology Bulletin, Series 14]*, 48*(3), 294–317. https://doi.org/10.11621/LPJ-25-34

OECD. (2021). *Education at a glance 2021: OECD indicators.* OECD Publishing. https://doi.org/10.1787/b35a14e5-en

Page, M.J., McKenzie, J.E., Bossuyt, P.M., Boutron, I., Hoffmann, T.C., Mulrow, C.D., Shamseer, L., Tetzlaff, J.M., Akl, E.A., Brennan, S.E., Chou, R., Gluud, C., Grimshaw, J.M., Hróbjartsson, A., Lalu, M.M., Li, T., Loder, E.W., Mayo-Wilson, E., McDonald, S., … Higgins, J.P.T. (2021). The PRISMA 2020 statement: An updated guideline for reporting systematic reviews. *BMJ, 372*, n71. https://doi.org/10.1136/bmj.n71

Paley, V.G. (1990). *The boy who would be a helicopter.* Harvard University Press.

Petitpierre, G., & Gyger, M. (2020). Early childhood education and care in Switzerland: Challenges and recent developments. In S. Garvis, S. Phillipson, & H. Harju-Luukkainen (Eds.), *International perspectives on early childhood education and care* (pp. 148–162). Routledge.

Popay, J., Roberts, H., Sowden, A., Petticrew, M., Arai, L., Rodgers, M., Britten, N., Roen, K., & Garner, P. (2006). *Guidance on the conduct of narrative synthesis in systematic reviews: A product from the ESRC methods programme* (Version 1). Lancaster University. https://doi.org/10.13140/2.1.1018.4643

Pramling Samuelsson, I., & Asplund Carlsson, M. (2008). The playing learning child: Towards a pedagogy of early childhood. *Scandinavian Journal of Educational Research, 52*(6), 623–641. https://doi.org/10.1080/00313830802497265

Pyle, A., DeLuca, C., & Danniels, E. (2017). A scoping review of research on play-based pedagogies in kindergarten education. *Review of Education, 5*(3), 311–351. https://doi.org/10.1002/rev3.3097

Rao, N., Sun, J., Chen, E.E., & Ip, P. (2021). Effectiveness of early childhood interventions in promoting cognitive development in developing countries: A systematic review and meta-analysis. *Hong Kong Journal of Paediatrics, 26*(1), 23–36.

Richard, S., Baud-Bovy, G., Clerc-Georgy, A., & Gentaz, E. (2021). The effects of a pretend play-based training designed to promote the development of emotion comprehension, emotion regulation, and prosocial behaviour in 5- to 6-year-old Swiss children. *British Journal of Psychology, 112*(3), 690–719. https://doi.org/10.1111/bjop.12484

Ryabkova, I.A., Sokolova, S.P., & Lvova, N.V. (2025). Igra ponaroshku: priniatie det'mi rannego vozrasta mnimoi situatsii. [Pretend play: Young children's acceptance of a make-believe situation]. *Natsionalnii Psikhologicheskii Zhurnal* [National Psychological Journal], *20*(4), 33–45.https://doi.org/10.11621/npj.2025.0403

Serafimovich, I.V., & Zakharova, T.N. (2025). Vzaimodeistvie kognitivnykh i lichnostnykh resursov pedagogov kak usloviye sovershenstvovaniya kachestva doshkol'nogo obrazovaniia [Interaction of cognitive and personal resources of preschool teachers as the condition of improving the quality of education]. *Sovremennoe doshkol'noe obrazovanie* [Preschool Education Today], *19*(4), 30–45. https://doi.org/10.24412/2782-4519-2025-4130-30-45

Skene, K., O'Farrelly, C.M., Byrne, E.M., Kirby, N., Stevens, E.C., & Ramchandani, P.G. (2022). Can guidance during play enhance children's learning and development in educational contexts? A systematic review and meta-analysis. *Child Development, 93*(4), 1162–1180. https://doi.org/10.1111/cdev.13730

Smilansky, S., & Shefatya, L. (1990). *Facilitating play: A medium for promoting cognitive, socio-emotional and academic development in young children.* Psychosocial and Educational Publishers.

Sterne, J.A. C., Hernán, M.A., Reeves, B.C., Savović, J., Berkman, N.D., Viswanathan, M., Henry, D., Altman, D.G., Ansari, M.T., Boutron, I., Carpenter, J.R., Chan, A.-W., Churchill, R., Deeks, J.J., Hróbjartsson, A., Kirkham, J., Jüni, P., Loke, Y.K., Pigott, T.D., … Higgins, J.P.T. (2016). ROBINS-I: A tool for assessing risk of bias in non-randomised studies of interventions. *BMJ, 355*, i4919. https://doi.org/10.1136/bmj.i4919

Sterne, J.A.C., Savović, J., Page, M.J., Elbers, R.G., Blencowe, N.S., Boutron, I., Cates, C.J., Cheng, H.-Y., Corbett, M.S., Eldridge, S.M., Hernán, M.A., Hopewell, S., Hróbjartsson, A., Junqueira, D.R., Jüni, P., Kirkham, J.J., Lasserson, T., Li, T., McAleenan, A., … Higgins, J. P. T. (2019). RoB 2: A revised tool for assessing risk of bias in randomised trials. *BMJ, 366,* l4898. https://doi.org/10.1136/bmj.l4898

Sun, J., Rao, N., & Yu, S. (2022). Early childhood education and care in China: Policies and trends in the context of 'two-child policy'. In I. Sussmuth & T. Soo-Hoo (Eds.), *The landscape of preschool education: Cross-national and cross-cultural perspectives* (pp. 112–128). Routledge.

Tománková, D., & Vávrová, S. (2021). Implementation of play-based pedagogies in Slovak early childhood education: Challenges and opportunities. *European Early Childhood Education Research Journal, 29*(3), 389–403. https://doi.org/10.1080/1350293X.2021.1928718

Utami, A.D., Fleer, M., & Li, L. (2023). The 'player' role of the teacher in playworld creates new conditions for children's learning and development. *International Journal of Early Childhood, 55*(1), 169–186. https://doi.org/10.1007/s13158-022-00333-y

Vartiainen, J., Sormunen, K., & Kangas, J. (2024). Relationality of play and playfulness in early childhood sustainability education. *Learning and Instruction, 93,* 101963.

Veresov, N.N., Veraksa, A.N., & Plotnikova, V.A. (2025). PlayWorld to promote narrative development: Evidence from a double-blind control experiment. *Psychology in Russia: State of the Art, 18*(3), 140–157. https://doi.org/10.11621/pir.2025.0309

Vidal Carulla, C., Christodoulakis, N., & Adbo, K. (2021). Development of preschool children's executive functions throughout a play-based learning approach that embeds science concepts. *International Journal of Environmental Research and Public Health, 18*(2), 588. https://doi.org/10.3390/ijerph18020588

Vygotsky, L.S. (2016). Play and its role in the mental development of the child. *International Research in Early Childhood Education, 7*(2), 3–25.

Weisberg, D.S., Hirsh-Pasek, K., & Golinkoff, R.M. (2013). Guided play: Where curricular goals meet a playful pedagogy. *Mind, Brain, and Education, 7*(2), 104-112. https://doi.org/10.1111/mbe.12015

Zhang, W., Li, L., & Disney, L. (2025). Exploring a Chinese kindergarten teacher's perezhivanie when teaching mathematics in an educational experiment. *Learning, Culture and Social Interaction, 54,* 100938. https://doi.org/10.1016/j.lcsi.2025.100938

Original manuscript received April 22, 2026
Revised manuscript accepted August 2, 2026
First published online September 1, 2026

To cite this article: Sukhikh, V.L., Veresov, N.N., Gavrilova, M.N. (2026). Teacher Implementation of Sociodramatic Play in Early Childhood Education: A Systematic Review. *Psychology in Russia: State of the Art*, *19*(3), 45–72. DOI: 10.11621/pir.2026.0303

Psychology in Russia: State of the Art
Volume 19, Issue 3, 2026

Russian Psychological Society
Lomonosov Moscow State University

COGNITIVE PSYCHOLOGY

Experimental Modelling of Body Schema with the Rubber Hand Illusion and the Hand Laterality Judgment Task: Cognitive and Meta-cognitive Levels

Olga R. Dobrushina[a], Galina A. Arina[b], Daria A. Kazantseva[b*], Mariia M. Tsypushtanova[a], Angelina G. Makarova[a], Larisa A. Dobrynina[a]

[a] *Research Center of Neurology, Moscow, Russia*
[b] *Lomonosov Moscow State University, Russia*

*Corresponding author. E-mail: da.al.kaz.00@gmail.com

Background. The concept of body schema continues to evolve, and the taxonomy of body representations remains debated. Two well-established paradigms provide experimental access to it: the rubber hand illusion (RHI), which modifies the body schema by incorporating a new object into it, and the hand laterality judgment task (HLJT), which requires participants to determine the laterality of a depicted hand and has been shown to involve mental movement.

Objective. To investigate body schema components using two well-validated experiments, the RHI and the HLJT.

Design. Thirty-one right-handed healthy volunteers aged 18–35 years completed the HLJT and then took part in the RHI. Proprioceptive drift and sense of ownership served as measures of RHI onset, whereas accuracy and response time were recorded in the HLJT. Data analysis was performed with mixed linear modelling.

Results. A cross-analysis of the performance in these tasks revealed distinct associations: 1) between higher accuracy in the HLJT and slower proprioceptive drift in the RHI; and 2) between longer response time in the HLJT and early emergence of the sense of ownership over an artificial hand in the RHI. We propose that the first pair of variables reflects the strength of the structural body representation (the cognitive level of body schema), whereas the second reflects confidence in those representations (the metacognitive level of body schema).

Conclusion. We observed that participants with better performance on the HLJT showed weaker RHI effects: stronger internal body representations were less susceptible to illusion-related artificial distortions.

Keywords: body schema, rubber hand illusion, hand laterality judgment task, predictive coding, somato-perception

© Dobrushina, O.R., Arina, G.A., Kazantseva, D.A., Tsypushtanova, M.M., Makarova, A.G., Dobrynina, L.A., 2026

ISSN 2074-6857 (Print) / ISSN 2307-2202 (Online)
http://psychologyinrussia.com

Public significance statements

Body schema — the internal representation of the body's shape and posture — is disrupted in a number of disorders that impair quality of life, including stroke, eating disorders and phantom pain. Based on the rubber hand illusion and hand laterality judgment task experiments, this study proposes that body schema representations be separated into cognitive and metacognitive levels. Differentiating the body schema concept may open new possibilities for developing diagnostic and therapeutic methods for disorders affecting body schema.

Introduction

The body schema concept, which was proposed more than a century ago (Head, 1920), is still evolving. Originally, "body schema" referred to a constantly updated representation of the body's shape and posture through the integration of proprioceptive signals. Later, it became clear that not only proprioception but also visual, sensorimotor, and other sensory inputs contribute to updating the body schema, and that this multisensory integration leads to a supramodal (amodal) representation of the body's spatial properties (Graziano & Botvinick, 1999). However, not all authors accept that the body schema is unified (Matsumiya, 2021). There is broad agreement that the main function of the body schema is action guidance, including real and mentally simulated actions (Jeannerod, 2001; de Vignemont, 2009), which distinguishes it from body image. Body image, by contrast, refers to body representations unrelated to movement, including perceptual, conceptual and emotional dimensions (Gallagher, 2005). The embodied cognition framework proposes a broader role: the body schema may shape high-level cognitive processes such as a sense of agency (Gallagher, 2005). A further recent advance is the predictive coding account, in which body representations are continuously updated through the interaction between a priori top-down predictions and the current bottom-up sensory input (Apps & Tsakiris, 2014; Di Lernia et al., 2023; Morasso & Mohan, 2020; Owens et al., 2018).

The ongoing theoretical development, influenced by experimental studies and clinical observations (Sattin et al., 2023; Schwoebel & Coslett, 2005), aims to differentiate between multiple body representations. While there is growing consensus regarding the distinction between body schema and body image (de Vignemont, 2009; Gallagher, 2005), some authors distinguish between body structural description and body semantics (Schwoebel & Coslett, 2005). In addition, Longo (2016) describes six body representations, which belong either to somatoperception (body image, body schema, superficial schema and body model) or to somatorepresentation (semantic domain and structural description). Within the body schema, de Vignemont distinguishes protective and working forms. The protective body schema provides a sense of ownership and includes the affective dimension, whereas the working body schema supports action in space and provides a sense of agency (de Vignemont, 2017). Thus, the taxonomy of bodily representations remains unsettled.

To support theoretical advancement through data-driven approaches, multiple methods of experimental body schema modelling have been developed. Body schema modelling refers to experimental tasks that engage the body schema. The present

study, which evaluates components of the body schema, is based on a cross-analysis of performance in two well-validated body schema experiments — the rubber hand illusion (RHI) (Botvinick & Cohen, 1998) and the hand laterality judgment task (HLJT) (Parsons, 1994). During the RHI, the participant's real hand is hidden from view while they observe an artificial hand placed congruently with the body. Synchronous strokes to the real and artificial hands elicit the illusion of ownership over the artificial hand. Illusion strength is usually measured by the drift of the perceived hand position towards the artificial hand (proprioceptive drift) and by questionnaires assessing the subjective feeling of ownership. These two variables correlate in some but not all experiments, and they are believed to represent different aspects of the illusion rather than being causally related (Abdulkarim & Ehrsson, 2016; Rohde et al., 2011). The RHI shows perceptual and neuronal stability across different settings and over time, so it is believed to reflect trait-like features of body perception (Bekrater-Bodmann et al., 2012).

The RHI is thought to depend on multisensory integration mechanisms, which recently have been conceptualized within the predictive coding framework: the strength of the illusion is believed to depend on the balance between bottom-up processing (visual and tactile stimulation) and top-down influences (the stability of body schema against distortions) (Dobrushina et al., 2021; Samad et al., 2015). Both bottom-up and top-down influences remain subjects of lively debate (Chancel et al., 2021; Ehrsson et al., 2022; Golaszewski et al., 2021; Limanowski, 2021; Lush et al., 2020; Rossi Sebastiano et al., 2022; Santos et al., 2023; Slater & Ehrsson, 2022; Thériault et al., 2022). Interest in the dynamic characteristics of the RHI and in its fading has grown in recent years (Abdulkarim et al., 2021; Finotti et al., 2023; Perepelkina et al., 2018). The present study used a dynamic version of the rubber hand illusion, as it allows more detailed quantification of the processes related to predictive coding. Over the course of stimulation, illusion strength increased gradually, consistent with progressive resolution of visuo-proprioceptive prediction error as the hypothesis that the artificial hand is one's own came to provide the best explanation of the incoming visuo-tactile input (Perepelkina et al., 2018).

The HLJT is a widely used body schema task that requires participants to distinguish left from right hands presented in different positions (Kim & Yi, 2021; Parsons, 1987, 1994; Sattin et al., 2023). Participants typically perform the task by mentally rotating their own hand into the presented position, which implies reliance on the body schema. This strategy is prevalent in children as early as 5–8 years of age (Spruijt et al., 2015). Moreover, limited experience of hand use impairs performance (Funk & Brugger, 2007). However, alternative accounts dispute the necessity of mental movement (Vannuscorps et al., 2012; Viswanathan et al., 2012). The hand-binding hypothesis explains the HLJT via a cross-modal analysis of the features of a seen hand, whereas mental movement is considered an aftereffect of this feature-based analysis (Viswanathan et al., 2012). Other authors regard mental movement as only one of several possible strategies for completing the HLJT (Moreno-Verdú & Hardwick, 2022). Performance is measured by the number of correct responses and by response time, which is longer for the images of hands rotated into uncommon, uncomfortable positions (Parsons, 1994).

To our knowledge, the RHI and HLJT have not been used simultaneously in previous studies. The RHI effect is thought to reflect the incorporation of an object into

the protective body schema (de Vignemont, 2017), whereas the mental movement involved in HLJT refers to the working body schema (Haslam et al., 2022). Thus, we aimed to map the links between the different components of the body schema using the RHI (proprioceptive drift and ownership) and the HLJT (answer correctness and response time). We considered a stronger RHI to be evidence of a less stable body schema, and successful HLJT completion to be evidence of stronger body representations. We hypothesized that higher scores in the RHI would be associated with lower scores in the HLJT.

Methods

Participants

The study sample included 31 (17 male) right-handed healthy volunteers aged 18–35 years (mean 20.5 ± 3.0). None of the participants had previously taken part in experiments with bodily illusions. Each participant completed two successive experiments — the HLJT followed by the RHI. The order of tasks was not balanced. The HLJT data were collected with the participation of N.V. Engovatova and were partially used in her master's thesis (Engovatova, 2022).

Procedure

Rubber Hand Illusion

We used a previously validated dynamic modification of the RHI (Perepelkina et al., 2018). The participants sat still at a table with their right hand hidden inside a wooden box. The middle finger of the right hand was positioned 20 cm from the edge of the table. A life-sized and gender-matched right-hand prosthesis ("rubber hand") was placed medial to the real right hand at a distance of 15 cm so that the participant could see it. The right forearm and the forearm of the rubber hand were covered with black fabric (*Figure 1*).

Figure 1. Experimental setting in the rubber hand illusion

The experiment consisted of two conditions: the emergence of the illusion and the fading of the illusion. During the emergence condition, the experimenter stroked the real and artificial hands slightly irregularly at a mean frequency of 0.5 Hz. During the fading condition, stimulation was absent: the participant could no longer see the rubber hand (the experimenter had removed it), and there was no tactile stimulation. We used a behavioral measure (proprioceptive drift) and a subjective measure (ownership questionnaire) of illusion strength. Proprioceptive drift was measured as the difference between the initial error in right-hand localization (measured before the start of the stimulation) and the perceived position of the hand after 15, 30, 60, 120 and 240 seconds from the onset of each condition. With their eyes closed, participants made a pointing movement to the middle finger of the right hand from each of three starting points, presented in randomized order. The results of these three measurements were averaged. The ownership questionnaire was completed at the same time points only during the emergence condition. It consisted of three test questions (statements related to the real effects of RHI) and three control questions rated on a Likert scale from -3 to 3; the questionnaire was published previously (Perepelkina et al., 2019).

All participants completed the RHI after the computerized HLJT.

Hand Laterality Judgment Task

We used a computerized HLJT based on Parsons (1994). The stimulus material comprised black-and-white photographs of both hands (left and right) in palm-up and palm-down positions. These four photographs were digitally manipulated to create images of hands rotated to 0°, 90° medial, 180° and 90° lateral orientations, giving a total of 16 different images (*Figure 2*). Each image was presented four times in a

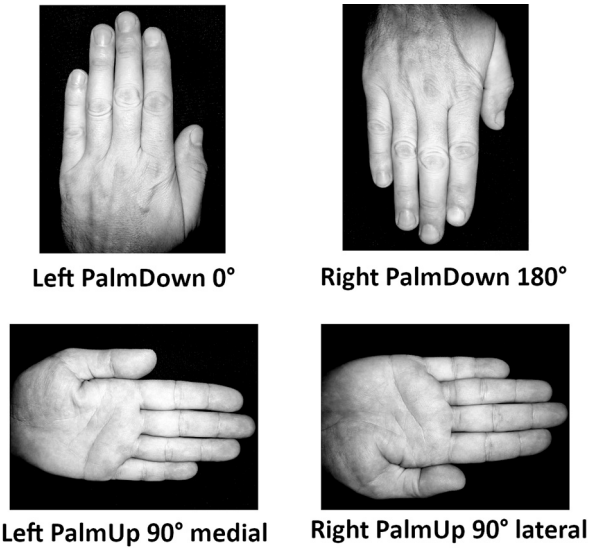

Figure 2. Examples of Stimuli in the Hand Laterality Judgment Task

Note. Hands differed in handedness (right or left), position (palm up or down) and orientation (0°, 90° medial, 180° and 90° lateral), resulting in 2 × 2 × 4 = 16 different stimuli.

sequence with a randomized stimulus order (64 trials in total). The presentation of the task and the collection of responses were performed using the Cogent Matlab Toolbox (http://www.vislab.ucl.ac.uk/cogent_2000.php, RRID:SCR_015672).

Participants sat at a computer with their left hand resting on the table and two fingers of their right hand on the response keys ("<" — left, ">" — right). Each stimulus appeared on the screen and remained there until the participant responded. Participants were asked to identify the laterality of the hand as quickly and accurately as possible. Before the test, instructions were displayed on the screen, and a practice stimulus was then presented.

Statistical Analysis

Data analysis was performed using mixed linear modelling in R Project 3.6.0 (https://www.R-project.org/, RRID:SCR_001905), packages "lme4", "lmerTest", "dplyr", "ggplot2", "simr". Continuous variables were rescaled.

First, we modelled the data from the two experiments separately in order to verify their validity. For the RHI, the following model was used for the proprioceptive drift dynamics:

Proprioceptive drift ~ Time * Condition + (1 | Participant) (1),

where "Time" represented the time of stimulation, "Condition" indicated the emergence or fading phase of the illusion and (1 | Participant) represented random participant-specific effects. For the ownership dynamics, the following model was used:

Ownership questionnaire ~ Time * Subscale + (1 | Participant) (2),

where "Time" represented the time of stimulation and "Subscale" indicated test or control questionnaire statements.

For the HLJT, the following models were used for the answer correctness and response times for each stimulus:

Answer correctness ~ Hand * Palm * Orientation + (1 | Participant) (3),

Log (Response time) ~ Hand * Palm * Orientation + (1 | Participant) (4),

where the stimulus characteristics "Hand" (right or left), "Palm" (up or down) and "Orientation" (0°, 90° medial, 180° or 90° lateral) served as independent variables. A binomial distribution was used to model answer correctness (0 or 1).

Spearman correlation was used for pairwise evaluation of the dependencies between the total number of correct responses and mean response time, and also between the ownership questionnaire score and proprioceptive drift.

Second, to assess the interaction between the performance in the two experiments, we entered the resulting parameters from the HLJT — the total number of correct responses (Accuracy) and the average correct response time (Response Time) — into the models representing RHI dynamics:

Proprioceptive drift ~ Time * Condition * Accuracy + (1 | Participant) (5)

Proprioceptive drift ~ Time * Condition * Average Response Time + (1 | Participant) (6)

Ownership questionnaire ~ Time * Subscale * Accuracy + (1 | Participant) (7)

Ownership questionnaire ~ Time * Subscale * Average Response Time + (1 | Participant) (8)

The significance threshold for these four confirmatory models was Bonferroni-corrected to $\alpha = .05/4 = .0125$.

The normality of the residual distributions for all linear mixed models was assessed by skewness and kurtosis in accordance with the reference values proposed by Kim (2013). All models were tested for homoscedasticity of the residuals. Where heteroscedasticity was detected, mixed models with separate residual variances for each of the predictors (fixed and random) were fitted (Nakagawa et al., 2025). Heterogeneity did not affect the results in any case.

Transparency and Openness

We report how we determined our sample size, all data exclusions, all manipulations, and all measures in the study. All data, code and stimuli materials are available in the Mendeley Repository at http://dx.doi.org/10.17632/p7zv55hnd5.1. This study's design and its analysis were not pre-registered.

The sample size was similar to those of previous studies utilizing the RHI (Perepelkina et al., 2017, 2018) and the HLJT (Parsons, 1987, 1994; Schwoebel, 2001). Power analysis was performed using Monte Carlo simulations and, in the absence of previous studies combining the HLJT and the RHI, was based on the observed effect sizes. For the model describing the dependence of proprioceptive drift on accuracy in the HLJT, the estimated power for the interactions of accuracy with time and with condition was over 99%. For the model describing the dependence of the ownership questionnaire on average response time in the HLJT, the estimated power for the interaction of average response time with the questionnaire subscale was 91%, and the power for the interaction of average response time, questionnaire subscale, and time was 73%.

Results

Rubber Hand Illusion

As expected, proprioceptive drift depended on time (estimate .08, 95% CI .05 to .11, $t = 5.29$, $p < .001$) and condition (estimate .04, 95% CI .02 to .06, higher during fading, $t = 3.62$, $p < .001$), with an interaction of time and condition (estimate –.09, 95% CI –.14 to –.05, negative effect of time during fading; $t = -4.16$, $p < .001$), along with a dependence of ownership on time (estimate .19, 95% CI .11 to .27, $t = 4.80$, $p < .001$) and subscale (estimate .47, 95% CI .41 to .52, higher score for test statements, $t = 16.06$, $p < .001$) with an interaction of time and subscale (estimate .13, 95% CI .02 to .24, higher effect of time for test statements, $t = 2.27$, $p = .02$). Thus, the dynamics of the emergence and fading phases were captured by both proprioceptive drift and ownership questionnaire responses.

The correlation between proprioceptive drift and the sum of test statements on the ownership questionnaire was not significant at any time point.

Hand Laterality Judgment Task

During the HLJT, errors were observed in 9.9% of trials. The total number of correct responses ranged from 41 to 64 (out of 64) with a median of 59 (25th quartile: 55;

75th quartile: 62). The error rate was higher for the 180° orientation (estimate –1.28, 95% CI –2.25 to –.31, $z = -2.60$, $p = .01$).

The average response time ranged from 1097 to 5965 ms with a median of 2101 ms (25th quartile 1529 ms, 75th quartile 2524 ms). Response time was longer for the palm-up position (estimate .21, 95% CI .10 to .31, $t = 3.82$, $p = .001$), 180° orientation (estimate .57, 95% CI .47 to .68, $t = 10.65$, $p < .001$) and 90° medial orientation (estimate .17, 95% CI .06 to .27, $t = 3.10$, $p = .002$), with an interaction of palm-up position and 180° orientation (estimate –.29, 95% CI –.43 to –.14, $t = -3.76$, $p = .01$) and also of palm-up position and 90° medial orientation (estimate –.23, 95% CI –.38 to –.08, $t = -3.02$, $p = .03$). Post hoc analyses were performed for the Orientation factor and the Orientation × Palm interaction (*Tables 1* and *2*).

Table 1

Post-hoc Analyses: Pairwise Comparisons by Orientation Factor

Orientation	Mean	SE	df	Lower CL	Upper CL	*p*-value (0)	*p*-value (180)	*p*-value (90lateral)
0	7.34	.06	34.60	7.22	7.47	–	–	–
180	7.78	.06	34.60	7.65	7.90	< .001	–	–
90lateral	7.57	.06	34.60	7.45	7.70	< .001	< .001	–
90medial	7.40	.06	34.60	7.27	7.52	1	< .001	< .001

Note. SE = Standard Error. CL = Confidence Limit.

Table 2

Post-hoc Analyses: Pairwise Comparisons by Orientation × Palm Interaction

Orientation	Palm	Mean	SE	df	Lower CL	Upper CL	*p*-value (0)	*p*-value (180)	*p*-value (90lateral)
0	Down	7.23	.07	41.30	7.09	7.36	–	–	–
180	Down	7.83	.07	41.30	7.70	7.96	< .001	–	–
90lateral	Down	7.40	.07	41.30	7.27	7.53	< .001	< .001	–
90medial	Down	7.42	.07	41.30	7.29	7.55	< .001	< .001	1
0	Up	7.46	.07	41.30	7.33	7.60	–	–	–
180	Up	7.73	.07	41.30	7.59	7.86	< .001	–	–
90lateral	Up	7.75	.07	41.30	7.61	7.88	< .001	1	–
90medial	Up	7.37	.07	41.30	7.24	7.51	.12	< .001	< .001

Note. SE = Standard Error. CL = Confidence Limit.

We observed the longest response times in the 180° orientation, intermediate response times in the 90° lateral orientation and the fastest response times in both the 90° medial and 0° orientations (*Table 1*). In the palm-down position, the longest

response time was observed in the 180° orientation, intermediate response times in the 90° medial and 90° lateral orientations, and the fastest response time in the 0° orientation. In the palm-up position, the longest response times were observed in the 180° and 90° lateral orientations, and the fastest in the 0° and 90° medial orientations (*Table 2*). Thus, more time was required for the mental rotation of the depicted hand into views that correspond to uncommon, uncomfortable positions of a real hand.

No correlation was observed between the total number of correct responses and the average response time (Rho = –.05, $p = .75$).

Accuracy in the Hand Laterality Judgment Task and Rubber Hand Illusion Dynamics

Accuracy in the mental hand rotation task was linked to the dynamics of proprioceptive drift, but not to the sense of ownership in the RHI. The influence of accuracy in the mental hand rotation task on proprioceptive drift dynamics is shown in the *Figure 3* (a median split is used for illustrative purposes). We observed an interaction of accuracy with time (estimate –.25, 95% CI –.37 to –.13, $t = -3.96$, $p < .001$), with condition (estimate –.24, 95% CI –.33 to –.15, $t = -5.38$, $p < .001$) and an interaction of the number of correct responses, condition and time (estimate .38, 95% CI .20 to .56, $t = 4.30$, $p < .001$). To distinguish the effects related to the emergence and fading phases, we performed a post hoc mixed-models analysis for these two phases separately. During the emergence phase, an interaction of accuracy with time was observed (high accuracy related to slower dynamics, estimate –.25, 95% CI –.37 to –.13,

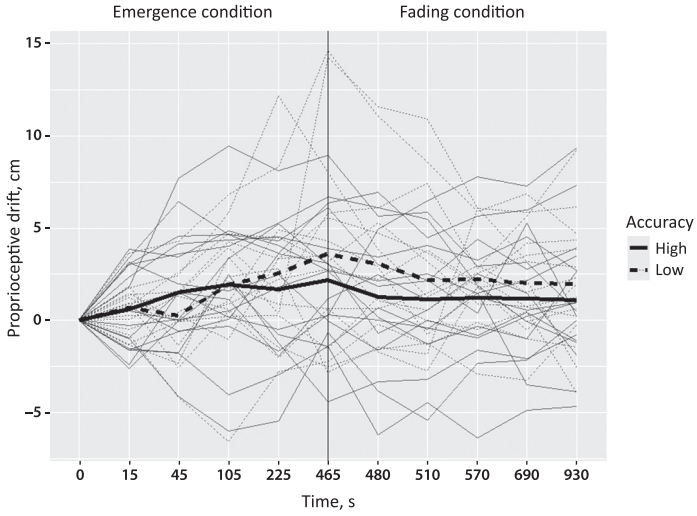

Figure 3. Interaction of the proprioceptive drift dynamics with the number of correct responses in the mental hand rotation task

Note. In the "spaghetti" plot, each line represents the dynamics of proprioceptive drift during the RHI in an individual participant. The time point 240 s corresponds to the end of the emergence phase (discontinuation of stimulation). For illustrative purposes, a median split of the sample was performed based on accuracy in the mental hand rotation task. Bold lines represent means of the two subsamples.

$t = -4.05$, $p < .001$), whereas during the fading phase higher accuracy was associated with a lower absolute value of drift (estimate –.18, 95% CI –.32 to –.04, $t = -2.75$, $p = .02$) and with less prominent dynamics (estimate .13, 95% CI .02 to .25, $t = 2.60$, $p = .02$). Taken together, higher accuracy in the mental hand rotation task was linked to less pronounced RHI dynamics.

Average Response Time in the Hand Laterality Judgment Task and Rubber Hand Illusion Dynamics

Response time in the mental hand rotation task was significantly associated with ownership dynamics, but not with proprioceptive drift in the RHI. For ownership, we observed an interaction between response time and question type (estimate .50, 95% CI .22 to .78, $t = 3.50$, $p < .001$) and between response time, question type, and time (estimate –.81, 95% CI –1.36 to –.25, $t = -2.81$, $p = .005$). To distinguish the effects related to the test and control questionnaire statements, we performed a separate post hoc mixed-models analysis. No effects were found for the control statements, whereas for the test statements a longer response time was associated with a higher absolute value (estimate .56, 95% CI .14 to .98, $t = 2.54$, $p = .01$) but slower dynamics (interaction with time: estimate –.57, 95% CI –.96 to –.17, $t = -2.82$, $p = .006$) of ownership. The dynamics of the ownership questionnaire (sum of test statements) as a function of response time are shown in the *Figure 4* (median split used for illustrative purposes). Both the modelling and visual inspection indicate that slow performance in the mental hand rotation task was associated with early emergence of the subjective feeling of ownership over the rubber hand.

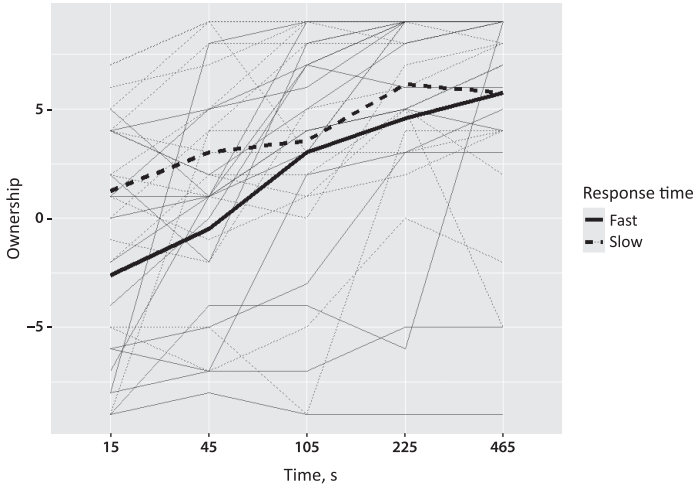

Figure 4. Interaction of the ownership questionnaire dynamics with response time in the mental hand rotation task

Note. In the "spaghetti" plot, each line represents the ownership questionnaire (sum of test statements) dynamics during the emergence phase of the RHI in an individual participant. For illustrative purposes, a median split of the sample was performed based on the response time in the mental hand rotation task. Bold lines represent means of the two subsamples.

Discussion

Our results are consistent with predictive coding accounts of body perception (Apps & Tsakiris, 2014; Morasso et al., 2015), of body ownership (Bertoni et al., 2023) and, in particular, of the RHI (Chancel et al., 2021; Dobrushina et al., 2021; Golaszewski et al., 2021; Samad et al., 2015; Santos et al., 2023). In the predictive coding framework, a stronger RHI indicates the predominance of bottom-up processes. The HLJT can also be conceptualized in terms of predictive coding, since the motor image serves as a prediction that is compared with the depicted hand before a decision is made. Participants with poor performance on the HLJT showed faster dynamics in the RHI experiment, which can be explained by less-developed body schema representations and, thus, weaker top-down influences that protect the body schema from artificial distortions. Results in the HLJT improve with motor practice (Habacha et al., 2014) and are negatively affected by hand immobilization (Debarnot et al., 2018), which suggests that the task reflects the strengths of hand representations. Thus, we propose that performance in the HLJT relies on the availability of bodily representations, *i.e.,* top-down processing. Both proprioceptive drift and the ownership questionnaire in the RHI were negatively associated with performance on the HLJT. These results parallel the absence of RHI effects in expert pianists (Pyasik et al., 2018), who, due to years of practice, may have highly developed internal representations of the hand.

A cross-analysis of the two experiments revealed distinct associations: first, between proprioceptive drift and accuracy in the HLJT, and second, between the sense of ownership over an artificial hand and response time in the HLJT. At the same time, we did not observe pairwise correlations between proprioceptive drift and ownership scores in the RHI or between accuracy and response time in the HLJT. Some RHI studies show an association between proprioceptive drift and ownership scores (Botvinick & Cohen, 1998; Longo et al., 2008; Samad et al., 2015; Tosi et al., 2023), whereas others fail to detect it (Holle et al., 2011; Rohde et al., 2011) and show that hand proprioception may change without changes in ownership (Holmes et al., 2006; Makin et al., 2008). The same dissociation between ownership and proprioceptive drift appears in somatic RHI (Gallagher et al., 2020). A study using subtle mechanical manipulation of participants' hand position during the RHI experiment found no causal link between proprioceptive drift and a sense of ownership (Abdulkarim & Ehrsson, 2016). According to a recent meta-analysis, these measures are related, but the correlation is rather weak (Tosi et al., 2023). The association between accuracy and response time in the HLJT remains largely underexplored. The limited available evidence comes from a few studies with differing designs, which reported a correlation between error rate and response time (Conson et al., 2012; Parsons, 1994).

We suggest that the two pairs of variables (proprioceptive drift and accuracy in the HLJT; ownership questionnaire and response time in the HLJT) characterize distinct domains of body schema organization. By analogy with other cognitive tasks, the two domains identified here may be defined as cognitive and metacognitive. In our interpretation, proprioceptive drift and accuracy in the HLJT reflect the spatial, cognitive level of the body schema, i.e., the strength of the structural representations. Participants with stronger body representations may be less susceptible to illusion-

related errors in localizing the hidden hand and therefore show lower proprioceptive drift. The sense of ownership and response time in the HLJT may characterize the subjective, metacognitive level of body schema-related processing. A faster response may indicate higher confidence in the answer (the average response time was calculated using only correct responses to minimize the possible influence of misclicks and guesses). Participants with high confidence in their body schema may be more inclined to rely on internal representations rather than on actual sensory input when they develop a sense of body ownership (dominance of top-down processing). Accordingly, they show lower scores on the ownership questionnaire. Nevertheless, the distinction between the cognitive and metacognitive components of the body schema requires careful verification in further research that controls for metacognitive variables.

Our findings may contribute to the theoretical development of the body schema concept, and they may also have clinical significance. The body schema is affected in a number of disorders that impair quality of life (de Vignemont, 2009). The rubber hand illusion and HLJT are used to characterize disease-related alterations of the body schema in patients with stroke (de Vries et al., 2013; Feenstra et al., 2016; Llorens et al., 2017; Haslam et al., 2022), eating disorders (Campione et al., 2017; Eshkevari et al., 2011), somatoform disorders (Perepelkina et al., 2019) and phantom pain (Giummarra et al., 2010). Theoretical models linking RHI and HLJT parameters to body schema dimensions can be used to interpret diagnostic data and develop targeted therapeutic interventions for patients with body schema distortions.

Conclusion

The results of our study support the predictive coding conceptualization of body schema, which is believed to be continuously updated through an interaction between prior inferences and current sensory input. We observed the expected negative association between RHI strength and HLJT performance. Beyond this, our results can be interpreted as a distinction between the cognitive and metacognitive levels of body schema organization. The cognitive level may correspond to structural representations, whereas the metacognitive level may comprise confidence in those representations. However, this hypothesis requires further confirmation. The proposed model now needs to be tested in further research using different experimental designs.

Limitations

The most substantial limitation of the study is the order of task presentation, which was not counterbalanced across participants. All participants first completed the HLJT and then took part in the RHI. Thus, the effect of presentation order was not controlled, which means that order, learning or carry-over effects cannot be excluded. For example, according to our hypotheses, we would expect reduced RHI scores after participation in the HLJT. Although the data show that, on average, the RHI was formed successfully, these measures may be underestimated because the preceding task also engaged the body schema. In addition, the small sample size limits the generalizability of the results.

Ethics Statement

The study protocol was approved by the Ethics Committee and the Institutional Review Board of the Research Center of Neurology (protocol number 4-3/21), and all participants gave informed consent.

Conflict of Interest

The authors declare no conflict of interest.

References

Abdulkarim, Z., & Ehrsson, H.H. (2015). No causal link between changes in hand position sense and feeling of limb ownership in the rubber hand illusion. *Attention Perception & Psychophysics, 78*(2), 707–720. https://doi.org/10.3758/s13414-015-1016-0

Abdulkarim, Z., Hayatou, Z., & Ehrsson, H.H. (2021). Sustained rubber hand illusion after the end of visuotactile stimulation with a similar time course for the reduction of subjective ownership and proprioceptive drift. *Experimental Brain Research, 239*(12), 3471–3486. https://doi.org/10.1007/s00221-021-06211-8

Apps, M.A., & Tsakiris, M. (2013). The free-energy self: A predictive coding account of self-recognition. *Neuroscience & Biobehavioral Reviews, 41*, 85–97. https://doi.org/10.1016/j.neubiorev.2013.01.029

Bekrater-Bodmann, R., Foell, J., Diers, M., & Flor, H. (2012). The perceptual and neuronal stability of the rubber hand illusion across contexts and over time. *Brain Research, 1452,* 130–139. https://doi.org/10.1016/j.brainres.2012.03.001

Bertoni, T., Mastria, G., Akulenko, N., Perrin, H., Zbinden, B., Bassolino, M., & Serino, A. (2023). The self and the Bayesian brain: Testing probabilistic models of body ownership through a self-localisation task. *Cortex, 167*, 247–272. https://doi.org/10.1016/j.cortex.2023.06.019

Botvinick, M., & Cohen, J. (1998). Rubber hands 'feel' touch that eyes see. *Nature, 391*(6669), 756. https://doi.org/10.1038/35784

Campione, G.C., Mansi, G., Fumagalli, A., Fumagalli, B., Sottocornola, S., Molteni, M., & Micali, N. (2017). Motor-based bodily self is selectively impaired in eating disorders. PLoS ONE, *12*(11), e0187342. https://doi.org/10.1371/journal.pone.0187342

Chancel, M., Hasenack, B., & Ehrsson, H.H. (2021). Integration of predictions and afferent signals in body ownership. *Cognition, 212,* 104722. https://doi.org/10.1016/j.cognition.2021.104722

Conson, M., Mazzarella, E., Donnarumma, C., & Trojano, L. (2012). Judging hand laterality from my or your point of view: Interactions between motor imagery and visual perspective. *Neuroscience Letters, 530*(1), 35–40. https://doi.org/10.1016/j.neulet.2012.09.051

De Vignemont, F. (2009). Body schema and body image — Pros and cons. *Neuropsychologia, 48*(3), 669–680. https://doi.org/10.1016/j.neuropsychologia.2009.09.022

De Vignemont, F. (2017). Agency and bodily ownership: the bodyguard hypothesis. *The Subject's Matter: Self-Consciousness and the Body*, 217. https://doi.org/10.7551/mitpress/10462.003.0013

De Vries, S., Tepper, M., Feenstra, W., Oosterveld, H., Boonstra, A.M., & Otten, B. (2013). Motor imagery ability in stroke patients: The relationship between implicit and explicit motor imagery measures. *Frontiers in Human Neuroscience, 7,* 790. https://doi.org/10.3389/fnhum.2013.00790

Debarnot, U., Huber, C., Guillot, A., & Schwartz, S. (2018). Sensorimotor representation and functional motor changes following short-term arm immobilisation. *Behavioral Neuroscience, 132*(6), 595–603. https://doi.org/10.1037/bne0000274

Di Lernia, D., Serino, S., Tuena, C., Cacciatore, C., Polli, N., & Riva, G. (2023). Mental health meets computational neuroscience: A predictive Bayesian account of the relationship between interoception and multisensory bodily illusions in anorexia nervosa. *International Journal of Clinical and Health Psychology, 23*(4), 100383. https://doi.org/10.1016/j.ijchp.2023.100383

Dobrushina, O.R., Arina, G.A., Dobrynina, L.A., Novikova, E.S., Gubanova, M.V., Belopasova, A.V., Vorobeva, V.P., Suslina, A.D., Pechenkova, E.V., Perepelkina, O.S., Kremneva, E.I., & Krotenkova, M.V. (2021). Sensory integration in interoception: Interplay between top-down and bottom-up processing. *Cortex, 144,* 185–197. https://doi.org/10.1016/j.cortex.2021.08.009

Ehrsson, H.H., Fotopoulou, A., Radziun, D., Longo, M.R., & Tsakiris, M. (2022). No specific relationship between hypnotic suggestibility and the rubber hand illusion. *Nature Communications, 13*(1), 564. https://doi.org/10.1038/s41467-020-18591-6

Ehrsson, H.H., Fotopoulou, A., Radziun, D., Longo, M.R., & Tsakiris, M. (2022). No specific relationship between hypnotic suggestibility and the rubber hand illusion. *Nature Communications, 13*(1), 564. https://doi.org/10.1038/s41467-022-28177-z

Engovatova, N.V. (2022). Individual'no-psikhologicheskie faktory resheniia zadachi po opredeleniiu lateralizatsii ruki. [Master's thesis, Lomonosov Moscow State University].

Eshkevari, E., Rieger, E., Longo, M.R., Haggard, P., & Treasure, J. (2011). Increased plasticity of the bodily self in eating disorders. *Psychological Medicine, 42*(4), 819–828. https://doi.org/10.1017/s0033291711002091

Feenstra, W., Tepper, M., Boonstra, A.M., Otten, B., & De Vries, S. (2016). Recovery of motor imagery ability in the first year after stroke. *International Journal of Rehabilitation Research, 39*(2), 171–175. https://doi.org/10.1097/MRR.0000000000000162

Finotti, G., Garofalo, S., Costantini, M., & Proffitt, D.R. (2023). Temporal dynamics of the Rubber Hand Illusion. *Scientific Reports, 13*(1), 7526. https://doi.org/10.1038/s41598-023-33747-2

Funk, M., & Brugger, P. (2007). Mental rotation of congenitally absent hands. *Journal of the International Neuropsychological Society, 14*(1), 81–89. https://doi.org/10.1017/S1355617708080041

Gallagher, M., Colzi, C., & Sedda, A. (2020). Dissociation of proprioceptive drift and feelings of ownership in the somatic rubber hand illusion. *Acta Psychologica, 212,* 103192. https://doi.org/10.1016/j.actpsy.2020.103192

Gallagher, S. (2005). *How the body shapes the mind.* Clarendon press.

Giummarra, M.J., Georgiou-Karistianis, N., Nicholls, M.E.R., Gibson, S.J., & Bradshaw, J.L. (2010). The phantom in the mirror: a modified rubber-hand illusion in amputees and normals. *Perception, 39*(1), 103–118. https://doi.org/10.1068/p6519

Golaszewski, S., Frey, V., Thomschewski, A., Sebastianelli, L., Versace, V., Saltuari, L., Trinka, E., & Nardone, R. (2021). Neural mechanisms underlying the Rubber Hand Illusion: A systematic review of related neurophysiological studies. *Brain and Behavior, 11*(8), e02124. https://doi.org/10.1002/brb3.2124

Graziano, M.S.A., & Botvinick, M.M. (2002). How the brain represents the body: Insights from neurophysiology and psychology. In *Common mechanisms in perception and action: Attention and performance XIX* (pp. 136–157). https://doi.org/10.1093/oso/9780198510697.003.0006

Habacha, H., Molinaro, C., Tabben, M., & Lejeune-Poutrain, L. (2014). Implementation of specific motor expertise during a mental rotation task of hands. *Experimental Brain Research, 232*(11), 3465–3473. https://doi.org/10.1007/s00221-014-4029-3

Haslam, B.S., Butler, D.S., Cocks, T.S., Kim, A.S., & Carey, L.M. (2022). Body schema as assessed by upper limb Left/Right judgment tasks is altered in stroke: Implications for motor imagery training. *Journal of Neurologic Physical Therapy, 47*(1), 26–34. https://doi.org/10.1097/NPT.0000000000000412

Head, H. (1920). *Studies in neurology,* Vol. 2. Oxford University Press.

Holle, H., McLatchie, N., Maurer, S., & Ward, J. (2011). Proprioceptive drift without illusions of ownership for rotated hands in the "rubber hand illusion" paradigm. *Cognitive Neuroscience, 2*(3–4), 171–178. https://doi.org/10.1080/17588928.2011.603828

Holmes, N.P., Snijders, H.J., & Spence, C. (2006). Reaching with alien limbs: Visual exposure to prosthetic hands in a mirror biases proprioception without accompanying illusions of ownership. *Perception & Psychophysics, 68*(4), 685–701. https://doi.org/10.3758/BF03208768

Jeannerod, M. (2001). Neural simulation of action: A unifying mechanism for motor cognition. *NeuroImage, 14*(1), S103–S109. https://doi.org/10.1006/nimg.2001.0832

Kim, A., & Yi, C. (2021). Understanding the Left Right Judgement Test: A literature review. *Physical Therapy Korea, 28*(4), 235–244. https://doi.org/10.12674/ptk.2021.28.4.235

Kim, H. (2013). Statistical notes for clinical researchers: assessing normal distribution (2) using skewness and kurtosis. *Restorative Dentistry & Endodontics, 38*(1), 52. https://doi.org/10.5395/rde.2013.38.1.52

Limanowski, J. (2021). Precision control for a flexible body representation. *Neuroscience & Biobehavioral Reviews, 134,* 104401. https://doi.org/10.1016/j.neubiorev.2021.10.023

Llorens, R., Borrego, A., Palomo, P., Cebolla, A., Noé, E., Badia, S.B.I., & Baños, R. (2017). Body schema plasticity after stroke: Subjective and neurophysiological correlates of the rubber hand illusion. *Neuropsychologia, 96,* 61–69. https://doi.org/10.1016/j.neuropsychologia.2017.01.007

Longo, M.R., Schüür, F., Kammers, M.P., Tsakiris, M., & Haggard, P. (2008). What is embodiment? A psychometric approach. *Cognition, 107*(3), 978–998. https://doi.org/10.1016/j.cognition.2007.12.004

Longo, M. (2016). Types of body representation. In Y. Coello & M.H. Fischer (Eds.), *Perceptual and emotional embodiment,* 1 (pp. 117–134). Routledge.

Lush, P., Botan, V., Scott, R.B., Seth, A.K., Ward, J., & Dienes, Z. (2020). Trait phenomenological control predicts experience of mirror synaesthesia and the rubber hand illusion. *Nature Communications, 11*(1), 4853. https://doi.org/10.1038/s41467-020-18591-6

Makin, T.R., Holmes, N.P., & Ehrsson, H.H. (2008). On the other hand: Dummy hands and peripersonal space. *Behavioural Brain Research, 191*(1), 1–10. https://doi.org/10.1016/j.bbr.2008.02.041

Matsumiya, K. (2022). Multiple representations of the body schema for the same body part. *Proceedings of the National Academy of Sciences, 119*(4). https://doi.org/10.1073/pnas.2112318119

Morasso, P., & Mohan, V. (2020). The body schema: Neural simulation for covert and overt actions of embodied cognitive agents. *Current Opinion in Physiology, 19*, 219–225. https://doi.org/10.1016/j.cophys.2020.11.009

Morasso, P., Casadio, M., Mohan, V., Rea, F., & Zenzeri, J. (2015). Revisiting the Body-Schema concept in the context of Whole-Body Postural-Focal dynamics. *Frontiers in Human Neuroscience, 9,* 83. https://doi.org/10.3389/fnhum.2015.00083

Moreno-Verdú, M., & Hardwick, R. (2022). Measuring the mind's eye: Does the Hand Laterality Judgement Task really assess motor imagery ability? PsyArXiv. https://doi.org/10.31234/osf.io/g2cj3

Nakagawa, S., Ortega, S., Gazzea, E., Lagisz, M., Lenz, A., Lundgren, E., & Mizuno, A. (2025). Location–scale models in ecology and evolution: Heteroscedasticity in continuous, count and proportion data. *Methods in Ecology and Evolution, 17*(2), 554–566. https://doi.org/10.1111/2041-210x.70203

Owens, A.P., Friston, K.J., Low, D.A., Mathias, C.J., & Critchley, H.D. (2018). Investigating the relationship between cardiac interoception and autonomic cardiac control using a predictive coding framework. *Autonomic Neuroscience, 210*, 65–71. https://doi.org/10.1016/j.autneu.2018.01.001

Parsons, L.M. (1987). Imagined spatial transformations of one's hands and feet. *Cognitive Psychology, 19*(2), 178–241. https://doi.org/10.1016/0010-0285(87)90011-9

Parsons, L.M. (1994). Temporal and kinematic properties of motor behavior reflected in mentally simulated action. *Journal of Experimental Psychology Human Perception & Performance, 20*(4), 709–730. https://doi.org/10.1037/0096-1523.20.4.709

Perepelkina, O., Boboleva, M., Arina, G., & Nikolaeva, V. (2017). Higher emotional intelligence is associated with a stronger rubber hand illusion. *Multisensory Research, 30*(7–8), 615–637. https://doi.org/10.1163/22134808-00002577

Perepelkina, O., Romanov, D., Arina, G., Volel, B., & Nikolaeva, V. (2019). Multisensory mechanisms of body perception in somatoform disorders. *Journal of Psychosomatic Research, 127,* 109837. https://doi.org/10.1016/j.jpsychores.2019.109837

Perepelkina, O., Vorobeva, V., Melnikova, O., Arina, G., & Nikolaeva, V. (2018). Artificial hand illusions dynamics: Onset and fading of static rubber and virtual moving hand illusions. *Consciousness and Cognition, 65*, 216–227. https://doi.org/10.1016/j.concog.2018.09.005

Pyasik, M., Salatino, A., & Pia, L. (2018). Do movements contribute to sense of body ownership? Rubber hand illusion in expert pianists. *Psychological Research, 83*(1), 185–195. https://doi.org/10.1007/s00426-018-1137-x

Rohde, M., Di Luca, M., & Ernst, M.O. (2011). The rubber hand illusion: Feeling of ownership and proprioceptive drift do not go hand in hand. *PLoS ONE, 6*(6), e21659. https://doi.org/10.1371/journal.pone.0021659

Rossi Sebastiano, A., Bruno, V., Ronga, I., Fossataro, C., Galigani, M., Neppi-Modona, M., & Garbarini, F. (2022). Diametrical modulation of tactile and visual perceptual thresholds during the rubber hand illusion: A predictive coding account. *Psychological Research, 86*(6), 1830–1846. https://doi.org/10.1007/s00426-021-01608-0

Samad, M., Chung, A.J., & Shams, L. (2015). Perception of body ownership is driven by Bayesian sensory inference. *PLoS ONE, 10*(2), e0117178. https://doi.org/10.1371/journal.pone.0117178

Santos, D., Agostinho, M., Treister, R., & Canaipa, R. (2023). Correlations between within-subject variability of pain intensity reports and rubber hand illusion proprioceptive drift. *Neuroscience Letters, 810,* 137319. https://doi.org/10.1016/j.neulet.2023.137319

Sattin, D., Parma, C., Lunetta, C., Zulueta, A., Lanzone, J., Giani, L., Vassallo, M., Picozzi, M., & Parati, E.A. (2023). An overview of the body schema and body image: Theoretical models, methodological settings and pitfalls for rehabilitation of persons with neurological disorders. *Brain Sciences, 13*(10), 1410. https://doi.org/10.3390/brainsci13101410

Schwoebel, J. (2001). Pain and the body schema: Evidence for peripheral effects on mental representations of movement. *Brain, 124*(10), 2098–2104. https://doi.org/10.1093/brain/124.10.2098

Schwoebel, J., & Coslett, H.B. (2005). Evidence for multiple, distinct representations of the human body. *Journal of Cognitive Neuroscience, 17*(4), 543–553. https://doi.org/10.1162/0898929053467587

Slater, M., & Ehrsson, H.H. (2022). Multisensory integration dominates hypnotisability and expectations in the rubber hand illusion. *Frontiers in Human Neuroscience, 16,* 834492. https://doi.org/10.3389/fnhum.2022.834492

Spruijt, S., Jongsma, M.L.A., Van Der Kamp, J., & Steenbergen, B. (2015). Predictive models to determine imagery strategies employed by children to judge hand laterality. *PLoS ONE, 10*(5), e0126568. https://doi.org/10.1371/journal.pone.0126568

Thériault, R., Landry, M., & Raz, A. (2022). The Rubber Hand Illusion: Top-down attention modulates embodiment. *Quarterly Journal of Experimental Psychology, 75*(11), 2129–2148. https://doi.org/10.1177/17470218221078858

Tosi, G., Mentesana, B., & Romano, D. (2023). The correlation between proprioceptive drift and subjective embodiment during the rubber hand illusion: A meta-analytic approach. *Quarterly Journal of Experimental Psychology, 76*(10), 2197–2207. https://doi.org/10.1177/17470218231156849

Vannuscorps, G., Pillon, A., & Andres, M. (2012). Effect of biomechanical constraints in the hand laterality judgment task: Where does it come from? *Frontiers in Human Neuroscience, 6*, 299. https://doi.org/10.3389/fnhum.2012.00299

Viswanathan, S., Fritz, C., & Grafton, S.T. (2012). Telling the right hand from the left hand. *Psychological Science, 23*(6), 598–607. https://doi.org/10.1177/0956797611429802

Original manuscript received February 15, 2026
Revised manuscript accepted August 28, 2026
First published online September 1, 2026

To cite this article: Dobrushina, O.R., Arina, G.A., Kazantseva, D.A., Tsypushtanova, M.M., Makarova, A.G., Dobrynina, L.A. (2026). Experimental Modelling of Body Schema with the Rubber Hand Illusion and the Hand Laterality Judgment Task: Cognitive and Meta-cognitive Levels. *Psychology in Russia: State of the Art*, *19*(3), 73–88. DOI: 10.11621/pir.2026.0304

Psychology in Russia: State of the Art
Volume 19, Issue 3, 2026

Russian Psychological Society
Lomonosov Moscow State University

EDUCATIONAL PSYCHOLOGY

Self-Monitoring Skills in Online Learning and the Role of Metacognitive Prompts: An Exploratory Study

Anastasia A. Koval[a*], Anna Y. Gorbunova[a], Andrey I. Podolsky[a], Aleksandra V. Getman[a]

[a]*HSE University, Moscow, Russia*

*Corresponding author. E-mail: aakoval@hse.ru

Background. Self-monitoring is critical for online learning, yet many students struggle due to cognitive load. Students with higher self-monitoring skills manage this demand better, but which load types they handle effectively remains unknown. Metacognitive prompts are a common support tool, but it is unclear whether they amplify or compensate for skill differences.

Objective. To examine (1) how self-monitoring skills relate to intrinsic (ICL), extraneous (ECL), and germane (GCL) load and performance, and (2) whether metacognitive prompts moderate these relationships.

Design. Participants completed an online reading task with prompts, followed by measures of self-monitoring skills, cognitive load, and a performance test. Linear regression examined relationships between self-monitoring skills, cognitive load, and performance, and whether use of prompts moderated them.

Results. After FDR correction, higher self-monitoring skills showed marginal associations with lower ICL ($b = -.48$, $p = .054$), lower ECL ($b = -.72$, $p = .054$), and higher performance ($b = .92$, $p = .054$), but were unrelated to GCL ($b = .09$, $p = .703$). Prompts significantly weakened the relationships between self-monitoring skills and ICL ($b = .99$, $p = .024$), ECL ($b = 1.65$, $p = .016$), and performance ($b = -2.34$, $p = .016$).

Conclusion. Higher self-monitoring skills may help manage cognitive load more effectively but are unrelated to GCL, suggesting that online learners employ self-monitoring differently than in traditional settings. Preliminary evidence also suggests that metacognitive prompts may compensate for rather than amplify skill differences. When prompts are provided, outcomes depend less on pre-existing skills

Keywords:
self-monitoring, cognitive load, online learning, self-regulated learning, metacognition, metacognitive prompts

© Koval, A.A., Gorbunova, A.Y., Podolsky, A.I., Getman, A.V., 2026

ISSN 2074-6857 (Print) / ISSN 2307-2202 (Online)
http://psychologyinrussia.com

Introduction

Online learning is increasingly prevalent in modern education, including higher education (Pikhart et al., 2023), yet it presents unique challenges for learners (Pham, 2022). Unlike traditional face-to-face instruction, online learning is characterized by spatial, temporal, and material flexibility (Barak et al., 2016; Basar et al., 2021). Students can choose where and when to study, access diverse multimedia resources, and review the materials multiple times (Barak et al., 2016; Basar et al., 2021). However, this flexibility places greater demands on learners, as the learning process becomes less structured (Sistek-Chandler, 2019). Online learning also involves diminished physical presence of instructors and peers, leading to reduced real-time feedback, limited communication, and lower levels of engagement (Akpen et al., 2024; Pikhart et al., 2023; Singh et al., 2024). From the perspective of Galperin's theory of the stage-by-stage formation of mental actions (Galperin, 1969), these challenges reflect the reduced availability of external support. In traditional classrooms, instructors naturally provide such support in the form of orienting bases — a system of cues and conditions that guide students and help them avoid inefficient trial-and-error strategies (Galperin, 1969). This orienting basis scaffolds students' mental actions and the learning process as a whole (Podolsky, 2023). In online environments, where these supports are largely absent, students must rely more heavily on their own resources to monitor their understanding and regulate their learning. This is why students are increasingly required to engage in self-regulation (de Bruin et al., 2020).

Self-regulation is a goal-directed process in which learners actively plan, monitor, and adapt their cognition, motivation, and behavior during learning (Pintrich, 2000; Winne & Hadwin, 1998; Zimmerman, 2000). In Zimmerman's model of self-regulated learning (SRL), this process is conceptualized as a cyclical sequence of forethought, performance, and self-reflection phases (Zimmerman, 2000). In the forethought phase, learners engage in goal setting, strategic planning, and motivational processes (*e.g.,* self-efficacy, task value). During this phase, students establish standards against which subsequent performance can be evaluated. In the performance phase, learners actively engage in the task while employing self-monitoring and self-control. Central to this phase is *self-monitoring*, a metacognitive process through which learners track their understanding, progress, and task performance in real time. Self-monitoring generates ongoing feedback that informs self-control processes, enabling learners to adjust strategies, allocate effort, or redirect attention as needed (Zimmerman, 2000). Finally, in the self-reflection phase, learners evaluate their performance outcomes through self-judgments (*e.g.,* self-evaluation, causal attribution) and self-reactions (*e.g.,* satisfaction, adaptive/defensive responses). These reflections feed forward into subsequent forethought processes. Thus, self-monitoring plays a central, integrative role in SRL by linking performance to control processes and connecting current learning episodes with future regulation cycles.

Self-monitoring is considered cognitively demanding, as students have to learn while simultaneously monitoring their own understanding (de Bruin et al., 2020; Seufert, 2018; Zimmerman & Moylan, 2009). This challenge is amplified in online learning environments by reduced feedback, absence of peer and instructor support, and screen-based presentation, all of which impair students' ability to accurately

monitor their understanding (Ackerman & Lauterman, 2012; Mou, 2023; Viberg et al., 2020; Wong et al., 2019).

Cognitive Load Theory (CLT; Sweller et al., 1998) provides a useful framework for understanding why self-monitoring is cognitively demanding. At its core, CLT assumes a fundamental distinction between limited working memory and virtually unlimited long-term memory. Working memory enables the conscious processing of new information but is severely constrained in both capacity and duration, allowing only a small number of elements to be processed simultaneously. In contrast, long-term memory stores organized knowledge structures, or schemas, which can be retrieved and processed as single units, thereby reducing cognitive demands during learning (Sweller et al., 2019). Learning, from a CLT perspective, involves the construction and automation of these schemas, enabling more efficient processing over time.

Within this framework, cognitive load arises from the demands placed on working memory during learning. CLT distinguishes three types of cognitive load: *intrinsic load (ICL)*, which depends on the complexity of the material and learners' prior knowledge; *extraneous load (ECL)*, which results from instructional design; and *germane load (GCL),* which reflects cognitive resources devoted to meaningful processing and schema construction (Sweller et al., 1998; Sweller et al., 2019).

Initially, GCL was conceptualized as an additional source of cognitive demand (Sweller et al., 1998). However, more recent work reconceptualizes GCL not as an additional component of total cognitive load, but as a mechanism for reallocating cognitive resources from unproductive to productive processing (Sweller et al., 2019). Despite this theoretical discussion, broad consensus has been reached that germane processing — the productive investment of working-memory resources in meaningful learning activities — is essential for learning (Kalyuga & Singh, 2015; Greenberg & Zheng, 2023). What is referred to as GCL in the present study is therefore best understood as an indicator of this productive cognitive effort, reflecting the extent to which learners engage in schema construction and meaningful processing. This operationalization is aligned with a substantial body of empirical research in which GCL has been successfully measured as a separable construct and its sensitivity to instructional manipulations has been demonstrated (Beege & Ploetzner, 2025; Costley et al., 2025; Gorbunova et al., 2026; Krieglstein et al., 2023). Treating GCL as a distinct type of load allows for a better understanding of how instructional interventions work, as GCL represents learning-relevant processing (Merkt & Bodemar, 2024).

From a CLT perspective, self-monitoring can be understood as an additional layer of cognitive processing that operates alongside primary learning activities. Prior research shows that self-monitoring may affect different types of cognitive load in distinct ways (Seufert, 2018). Self-monitoring may increase ECL by introducing additional regulatory processing demands, exert limited influence on ICL because task complexity remains unchanged, and enhance GCL by supporting learning-relevant cognitive activities (Seufert, 2018).

Students with higher self-monitoring skills are better at managing cognitive load and demonstrate more effective self-monitoring during learning (Glogger-Frey et al., 2016; Schunk, 1982; Seufert, 2018; Sidi & Ackerman, 2024; Zimmerman & Schunk,

2011). Specifically, students with high self-monitoring skills tend to acquire, organize and use information more efficiently, enabling them to regulate their study time effectively, allocate cognitive resources more strategically and persist in the face of difficulty (Shi & Qu, 2022). In contrast, students with low self-monitoring skills show weaker regulatory control, resulting in poorly planned and less purposeful cognitive activity, even when their cognitive abilities are relatively high (Shi & Qu, 2022). From a resource perspective, these differences reflect that higher self-monitoring skills provide learners with greater metacognitive resources for managing cognitive demands and optimizing learning outcomes (Karpov et al., 2017). It is also worth noting that according to CLT, fostering self-monitoring skills through instructional interventions may be inherently limited (Sweller et al., 2019). However, empirical research suggests that self-monitoring skills can be developed through extended instructional interventions, typically requiring several weeks or months of sustained practice (*e.g.,* Dignath et al., 2023; Huff & Nietfeld, 2009; Zimmerman & Schunk, 2011).

Beyond cognitive load, self-monitoring skills are also consistently associated with positive learning outcomes (*e.g.,* Jung et al., 2024; Schraw et al., 2006; Zimmerman & Schunk, 2011). However, the strength of their relationship with academic performance is less consistent, with meta-analytic evidence suggesting effects that range from weak to moderate (Credé & Phillips, 2011). This variability likely reflects multiple contextual and methodological influences, indicating that further investigation could help refine our understanding of these relationships.

To address the challenges students face with self-monitoring, instructional tools can be designed to activate self-monitoring during learning (Zimmerman, 2000). In online learning environments, metacognitive prompts are widely used (Guo, 2022). They are specifically designed to support self-monitoring by directing students' attention to their understanding, progress, or strategy use (Bannert & Mengelkamp, 2013; Berthold et al., 2007; Müller & Seufert, 2018). They function by signaling when and how students should monitor themselves, thereby activating monitoring processes that might otherwise not be initiated spontaneously (Berthold et al., 2007; Müller & Seufert, 2018). Compared to more elaborate self-monitoring tools, metacognitive prompts are considered simple tools because they require minimal effort and do not necessitate real-time instructor involvement (Engelmann et al., 2021).

Prior research shows that self-monitoring support tools, including metacognitive prompts, improve students' self-monitoring skills (Dignath et al., 2023) and, when gradually withdrawn, function as metacognitive scaffolding (Azevedo & Hadwin, 2005). However, the interaction between students' existing self-monitoring skills and metacognitive prompts remains underexplored. From a theoretical perspective, students with stronger self-monitoring skills would be expected to use prompts more efficiently (Nelson & Narens, 1990; Zimmerman, 2000), as their higher skills enable them to recognize the relevance of prompts and translate monitoring information into effective regulatory actions. Moreover, the academic benefits of any self-monitoring support depend not only on prompts but on students' capacity for reflective processing and strategic adjustment (Guo, 2022). These considerations suggest that metacognitive prompts may be particularly beneficial for students with stronger self-monitoring skills, facilitating more efficient management of cognitive load and improved academic performance.

On the other hand, an alternative theoretical framework suggests a different prediction. According to Galperin's theory of the stage-by-stage formation of mental actions (Galperin, 1969), cognitive processes, including self-monitoring, develop through the internalization of external, materialized actions. From this perspective, metacognitive prompts can be understood as a form of external scaffolding that provides learners with an orienting basis for action. As learners repeatedly engage with such prompts, the externally guided monitoring processes may gradually become internalized, transforming into autonomous self-monitoring skills (Talyzina, 1981). This view implies that prompts could be particularly beneficial for students with weaker self-monitoring skills, as they offer structured support that compensates for underdeveloped internal resources. The resource approach to metacognition (Karpov et al., 2017) further extends this idea. According to this view, learners with lower self-monitoring skills possess fewer internal metacognitive resources, making them more dependent on external support. Metacognitive prompts may therefore serve as a substitutional resource — temporarily compensating for deficient internal resources — thereby reducing performance gaps between high- and low-skilled students. Thus, contrary to the amplification hypothesis, these theoretical perspectives predict that metacognitive prompts may function as a compensatory mechanism, benefiting weaker students more than stronger ones.

Overall, despite the theoretical and practical importance of self-monitoring skills in online learning, several gaps remain. First, it is unclear how self-monitoring skills relate to different types of cognitive load and academic performance in online learning environments. Second, while metacognitive prompts are widely used to support self-monitoring in online learning, it remains unknown whether such prompts amplify existing skill differences (benefiting stronger students) or compensate for them (benefiting weaker students). These competing predictions can be operationalized as distinct interaction patterns. The amplification hypothesis would be supported by a positive interaction between self-monitoring skills and prompt condition, indicating that prompts strengthen the relationship between self-monitoring skills and outcomes, thereby benefiting higher-skilled students more. Conversely, the compensatory hypothesis would be supported by a negative interaction, indicating that prompts weaken the relationship between self-monitoring skills and outcomes, thereby reducing the advantage of higher-skilled students and benefiting lower-skilled students more. To address these gaps, the present study investigated two research questions.

RQ1: How are students' self-monitoring skills related to different types of cognitive load and academic performance in online learning?

RQ2: How do metacognitive prompts moderate the relationship between students' self-monitoring skills, cognitive load, and academic performance?

Methods

Participants and Procedure

The sample comprised 40 Russian undergraduate students (77.5% female, age range 19-23 years, median = 20) enrolled in an online minor course. The students repre-

sented diverse academic backgrounds across the humanities, social sciences, and technical fields.

Participants completed a 40-minute asynchronous online self-study task. They studied a text on questionnaire theory (Rogozin & Ipatova, 2020) and prepared for a subsequent test. Each participant received an individual copy of the text in Google Docs and was permitted to annotate the material. Participants were randomly assigned to either a prompt condition ($n = 20$) or a control condition ($n = 20$). Participants in the prompt condition received metacognitive prompts embedded after each of the six substantive sections of the text. Adapted from Berthold et al. (2007), these prompts encouraged learners to evaluate their understanding and connect the content to prior knowledge by asking three questions: (1) "Which two or three key ideas from this text do I understand best?"; (2) "Overall, how confident am I in my understanding of this text? (0–100%)"; and (3) "To what extent was this content already familiar to me? (very familiar / partly familiar / not familiar)", followed by a brief prompt to reflect on prior knowledge: "If I already knew something about this topic, what prior knowledge or experiences did I draw on?". These prompts were designed to support self-monitoring during asynchronous online study.

After doing the learning task, students had 20-25 minutes to complete a questionnaire and a test. Following that, the participants took part in a debriefing session. The study was approved by the institutional ethics committee, and the informed consent was obtained from all participants.

Measures

Different types of cognitive load (ICL, ECL, GCL) were measured using a validated *15-item questionnaire developed by Krieglstein et al. (2023)*. Responses were recorded on a 9-point Likert scale ranging from 1 (strongly disagree) to 9 (strongly agree). Cronbach's alpha coefficients indicated good internal consistency for ICL ($\alpha = .74$), ECL ($\alpha = .88$), and GCL ($\alpha = .86$). Cognitive load was measured immediately after the learning task and before the test to capture the cognitive demands of the learning activity rather than the assessment.

Self-monitoring skills were measured using items from the *Comprehension Monitoring subscale of the Metacognitive Awareness Inventory* (MAI; Schraw & Dennison, 1994), in the validated Russian adaptation (Perikova & Byzova, 2022). Responses were recorded on a 5-point Likert scale ranging from 1 (totally disagree) to 5 (absolutely agree). Cronbach's alpha also indicated good reliability ($\alpha = .71$). Self-monitoring skills were measured after the learning task and the test. This construct was treated as a stable individual characteristic that is unlikely to change as a result of a short experimental intervention. This timing is consistent with recommendations in SRL research to prevent SRL measurement tools from inadvertently promoting any self-regulation, including self-monitoring (Dörrenbächer-Ulrich et al., 2021). Administering the questionnaire prior to the learning task was avoided to prevent artificially heightened self-monitoring in the control group.

Academic performance was measured using a *10-item multiple-choice comprehension test*, with each question having one correct answer. It assessed students' acquisition and retention of key concepts from the learning material, thereby capturing their immediate learning outcome. When carefully constructed, multiple-choice

tests provide a reliable and valid basis for inferences about academic performance (Brown & Abdulnabi, 2017).

Analytical Strategy

The analysis was done in R. Linear regression analyses were conducted to examine the relationships between students' self-monitoring skills and cognitive load, as well as academic performance (RQ1). To test whether metacognitive prompts have a potential to moderate the relationships between self-monitoring skills and the outcome variables, multiple linear regression models including interaction terms were estimated (RQ2). Prompt condition was coded as a dichotomous variable, with 0 representing the control condition (no prompts) and 1 representing the prompt condition. This coding was used in all regression analyses for RQ2 to examine the main effect of prompts and their interaction with self-monitoring skills. The interaction term served as a test of whether prompts amplify or compensate for individual differences in self-monitoring skills. To control for multiple comparisons, False Discovery Rate (FDR) correction with a threshold of $q < .05$ was used (Benjamini & Hochberg, 1995). This approach was chosen over more conservative corrections (*e.g.,* Bonferroni) to balance Type I error control with statistical power (Nakagawa, 2004).

Prior to regression analyses, Spearman's correlation was used to check for multicollinearity among predictors and to describe the relationships among all study variables. Spearman's correlation was selected because all variables were measured on ordinal Likert-type scales, which do not meet the interval-level assumptions required for Pearson correlation. No correlations exceeded the recommended threshold of $r > .90$ (Tabachnick & Fidell, 2006), indicating no multicollinearity concerns. Other regression assumptions were also evaluated, and no other substantial violations were detected. Although the sample size was modest, it was adequate for the estimated models given the limited number of predictors, and prior evidence that linear regression yields reliable coefficient estimates in small samples when assumptions are met (Austin & Steyerberg, 2015; Harrell, 2015).

Results

Descriptive statistics *(see Table 1)* and correlation analysis (*see Table 2*) for all variables are presented.

Table 1

Descriptive Statistics

	M	Median	SD	Min	Max	Skewness	Kurtosis
Self–monitoring skills	3.69	3.75	.65	2.33	5.0	–.29	–.45
ICL	3.36	3.20	.91	1.80	5.6	.43	–.49
ECL	3.70	3.80	1.39	1.20	6.2	.01	–1.00
GCL	4.92	5.20	.99	1.40	6.4	–1.78	3.47
Academic performance	7.20	7.00	1.83	4.00	10.0	.08	–1.01

Table 2

Correlation Analysis (Spearman)

	Self-monitoring skills	ICL	ECL	GCL	Academic performance
Self-monitoring skills	–				
ICL	–.445**	–			
ECL	–.263*	.464**	–		
GCL	.238	–.336*	–.386*	–	
Academic performance	.342*	–.281	–.158	–.289	–

*Notes. *p < .05, **p < .01*

RQ1. How are students' self-monitoring skills related to different types of cognitive load and academic performance in online learning?

After applying FDR correction for multiple comparisons (*see Table 3*), self-monitoring skills showed marginal associations with ICL (b = –.48, adjusted $p = .054$), ECL (b = –.72, adjusted $p = .054$), and academic performance (b = .92, adjusted $p = .054$). Self-monitoring skills were unrelated to GCL (b = .09, adjusted $p = .703$).

Table 3

Linear Regression Models Results for the Relationships between Self-Monitoring Skills, Cognitive Load (ICL, ECL, GCL), and Academic Performance

	ICL	ECL	GCL	Academic performance
Intercept	5.12 (<.001***)	6.37 (<.001***)	4.56 (<.001***)	3.81 (.025*)
Self-monitoring skills	–.48 (.033*)	–.72 (.033*)	.09 (.703)	.92 (.041*)
FDR-adjusted p-value for self-monitoring skills	.054†	.054†	.703	.054†
Adjusted R^2	.09	.09	.003	.08
Partial R^2	.11	.11	.004	.11

*Notes. † p < .10 (trend), *p < .05, ***p < .001; values in parentheses are original p-values*

RQ2: How do metacognitive prompts moderate the relationship between students' self-monitoring skills, cognitive load, and academic performance?

After applying FDR correction for multiple comparisons (*see Table 4*), significant interactions were found between self-monitoring skills and metacognitive prompt use for ICL (b = .99, adjusted $p = .024$), ECL (b = 1.65, adjusted $p = .016$), and academic performance (b = –2.34, adjusted $p = .016$). The overall model for GCL was not sig-

nificant ($p = .103$). Given this, and consistent with recommendations for interpreting interaction effects (Aiken & West, 1991), we do not interpret the GCL interaction term and exclude it from further discussion.

To assess the robustness of the interaction effects for ICL, ECL, and academic performance, bias-corrected and accelerated (BCa) bootstrap confidence intervals based on 5,000 resamples were computed (Efron & Tibshirani, 1994). The 95% BCa bootstrap confidence intervals supported all three interaction effects: ICL (CI [.163, 1.878]), ECL (CI [.828, 2.765]), and academic performance (CI [-3.867, -.965]). As none of the confidence intervals included zero, the interaction effects appeared robust to resampling.

A significant positive main effect of prompt condition on ECL was also found ($b = .99$, $p < .001$). This main effect should not be interpreted in isolation, as it reflects the average across all participants and is qualified by the significant interaction, which shows that prompts were beneficial primarily for students with lower self-monitoring skills.

Table 4

Multiple Linear Regression Models Results for the Interactions Between Self-Monitoring Skills and Metacognitive Prompt Use Predicting Cognitive Load and Academic Performance

	ICL	ECL	GCL	Academic performance
Intercept	3.16 (<.001***)	3.16 (<.001***)	5.09 (<.001***)	7.42 (<.001***)
Self-monitoring skills	–1.11(.002 **)	–1.8(<.001***)	.82(.04 *)	2.39(<.001***)
Metacognitive prompt use (prompt condition)	.34(.19)	.99(< .001**)	–.29(.33)	–.32(.53)
Interaction term	.99(.024*)	1.65(.008**)	–1.15(.024*)	–2.34(.008**)
FDR-adjusted p-value for the interaction term	.024*	.016*	.024*	.016*
BCa CI for interaction term	[.163, 1.878]	[.828, 2.765]	–	[–3.867, –.965]
Model F	4.24(.012*)	7.32(< .001**)	2.22(.103)	4.4(.009**)
Adjusted R^2	.19	.33	.09	.21
Partial R^2	.26	.38	.16	.27

*Notes. *p < .05, **p < .01, ***p < .001; interaction terms represent the Self-Monitoring skills × Metacognitive prompt use effect; values in parentheses are original p-values*

Simple slope analyses (*see Table 5*) indicated that in the control group, higher self-monitoring skills were associated with lower ICL ($b = –1.11$, 95% CI [–1.78, –.44]), lower ECL ($b = –1.81$, 95% CI [–2.74, –.88]), and higher academic performance ($b = 2.39$, 95% CI [1.06, 3.72]). In contrast, in the prompt group, these relationships were no longer statistically significant for ICL ($b = –.12$, 95% CI [–.64, .40]), ECL ($b = –.16$, 95% CI [-.89, .57]), or academic performance ($b = .05$, 95% CI [-.99,

1.09]). Consistent with the significant interaction effects, these results indicate that metacognitive prompts weakened the association between students' self-monitoring skills and these learning outcomes, reducing the extent to which performance and cognitive load depended on students' initial self-monitoring skills.

Table 5

Simple Slopes for the Relationship between Self-Monitoring Skills and Learning Outcomes by Condition

	Group	b	SE	95% CI
ICL	Control	–1.11	.33	[–1.78, –.44]
	Prompt	–.12	.26	[–.64, .40]
ECL	Control	–1.81	.46	[–2.74, –.88]
	Prompt	–.16	.36	[–.89, .57]
Academic performance	Control	2.39	.66	[1.06, 3.72]
	Prompt	.05	.51	[–.99, 1.09]

Discussion

The present study investigated how self-monitoring skills are related to different types of cognitive load and academic performance in online learning environments (RQ1), and whether metacognitive prompts, the most widely used self-monitoring support in online settings, moderate these relationships (RQ2).

For RQ1, after FDR correction, higher self-monitoring skills showed marginal associations with lower ECL and higher academic performance, consistent with prior work linking self-monitoring to more effective self-regulated learning (Schraw et al., 2006; Zimmerman, 2000; Zimmerman & Schunk, 2011). This pattern is theoretically consistent with the assumption that better self-monitoring helps students detect misunderstandings earlier and avoid inefficient strategies, freeing resources that would otherwise be spent on extraneous processing (Seufert, 2018). Two findings diverged from our expectations: self-monitoring skills were also marginally associated with lower ICL and were unrelated to GCL. The trend for a negative association with ICL may reflect differences in perceived task difficulty — higher-skilled learners may organize and "chunk" information more effectively, making the material feel less complex even when task demands are unchanged (Löffler et al., 2016). The null effect for GCL contrasts with prior findings suggesting that effective self-monitoring should increase GCL by promoting deeper engagement with the material (Dignath et al., 2023; Guo, 2022; Seufert, 2018). This discrepancy may be explained by a cost-benefit perspective on cognitive load in online learning, which states that in this environment students are more prone to invest additional cognitive effort in deeper processing only when the perceived benefits outweigh the associated cognitive costs (Skulmowski & Xu, 2021). In the present study, students with higher

self-monitoring skills may have primarily used these skills to reduce ICL and ECL, while the learning context provided insufficient incentive or opportunity to invest additional cognitive effort in elaborative processing, leading to a null effect for GCL (Gorbunova et al., 2026).

For RQ2, metacognitive prompts were associated with moderated relationships between self-monitoring skills, ICL, ECL, and academic performance. Contrary to the assumption that prompts would strengthen these relationships, the observed pattern was compensatory. The associations between self-monitoring skills and learning outcomes were weaker when metacognitive prompts were used, potentially making successful learning less dependent on students' initial self-monitoring skills. From a theoretical perspective, prompts may externalize the metacognitive process, thus reducing differences between high- and low-skilled students (Galperin, 1969; Karpov et al., 2017). Prompts may have functioned as external scaffolds that provided an orienting basis for action and filled internal resource deficits, benefiting weaker students more. At the same time, prompts may not have boosted the advantages of higher skills because they target basic self-monitoring, not the elaborative processing that more skilled learners might need. However, this interpretation is speculative, and it is also possible that prompts did not merely compensate for skill deficits but may have interfered with the learning processes of higher-skilled students. Specifically, the significant interaction suggests that for students with higher self-monitoring skills, the additional structure provided by prompts may have been redundant or even disruptive, potentially reducing their performance and increasing their perceived cognitive load. Thus, while prompts may support lower-skilled students, they may not be beneficial and could even be counterproductive for students who are already capable of effective self-monitoring.

Regarding the GCL interaction, it should not be interpreted as evidence of a reliable effect. Although the interaction term was significant in isolation, the overall model was not significant. At the same time, prior research indicates that increases in GCL typically require more demanding forms of self-monitoring support, such as learning journals or generative tasks, than the brief comprehension checks used in this study (Guo, 2022).

Conclusion

This exploratory study examined self-monitoring skills in online learning, their relationship with different types of cognitive load (ICL, ECL, GCL) and academic performance, and whether metacognitive prompts have the potential to moderate these relationships. After FDR correction, self-monitoring skills showed marginal associations with lower ICL and ECL and better performance, while the relationship with GCL in this study was insignificant. Metacognitive prompts appeared to act as a compensatory mechanism rather than an amplifier of existing skills: when prompts were used, students' outcomes tended to depend less on their pre-existing self-monitoring skills, potentially reducing skill-based differences in cognitive load and performance. However, these findings are preliminary and require replication with larger samples.

The findings suggest several tentative implications for instructional design. First, given that self-monitoring skills showed marginal associations with lower cognitive load and better performance, fostering these skills through explicit instruction or practice opportunities may help students manage the cognitive demands of online learning more effectively. Second, because metacognitive prompts appear to reduce students' reliance on their existing self-monitoring skills, they may serve as an equity-oriented support for students with weaker self-monitoring skills. Finally, when the instructional goal is to promote deeper processing and improve academic performance, metacognitive prompts should be combined with more complex self-monitoring support (*e.g.*, learning journals or generative tasks) rather than used solely as brief comprehension checks.

Limitations

Several limitations should be considered when interpreting these findings. First, this study was exploratory in nature, and the results should be regarded as preliminary. The sample size was modest ($N = 40$), which limits statistical power. This may explain why the associations between self-monitoring skills and cognitive load and performance became marginal after FDR correction. According to Tabachnick and Fidell (2006), such a sample size provides sufficient power for detecting only large effects ($R^2 \geq .26$), while adequate detection of medium effects ($R^2 \approx .13$) typically requires larger samples. In linear regression models, the observed effects were not large (partial $R^2 = .11$), suggesting that marginal *p*-values after FDR correction reflect the limited sample size rather than the absence of meaningful relationships. The modest sample size also limited the stability of the interaction effects, with a minimally adequate cases-to-predictors ratio of 10:1 (Harrell, 2015). Bootstrap resampling (5,000 iterations) provided some reassurance; however, this does not substitute for adequate power. Replication with larger samples is needed to confirm the reliability of these findings.

In addition, the measurement of self-monitoring skills after the learning task may have introduced state-dependent influences, as scores could reflect not only stable individual differences, but also reactions to task difficulty and test performance. This timing, while consistent with common practice in SRL research to avoid reactivity effects (Dörrenbächer-Ulrich et al., 2021), nevertheless limits our ability to draw causal inferences. Furthermore, the assessment format, a short multiple-choice test, may have been insufficiently sensitive to differences in deeper understanding, potentially reducing observable academic performance effects even when learning processes differed. Future research should employ more diverse learning tasks and assessments that capture transfer or elaborative understanding.

Further studies could also compare different types of prompts, including AI-generated adaptive prompts tailored to individual learners' comprehension levels, and investigate how their use interacts with longer-term development of self-monitoring skills to better understand how compensatory support can transition into sustained self-regulated learning.

Ethics Statement

This study was conducted in accordance with the ethical standards of the Institutional Review Board of the National Research University Higher School of Economics (HSE IRB). The project titled "Supporting Self-Monitoring in Online Learning: Impact on Cognitive Load and Learning Outcomes" was reviewed and approved at the HSE IRB meeting held on October 23, 2025. The IRB confirmed that the research conforms fully to ethical guidelines and may be implemented in its current form.

All participants were informed about the purpose of the study, the procedures involved, and their right to withdraw at any time without penalty. Written informed consent was obtained from all participants prior to their involvement in the research. Confidentiality and anonymity of the data were ensured throughout the study.

Author Contributions

A.A. and A.Y. conceived of the idea. A.A. and A.V. collected the data. A.A. developed the theory and performed the computations. A.Y. and A.V. verified the analytical methods. A.Y. supervised the findings of this work. A.I. contributed to the conceptual explanation of the findings and assisted in the preparation of the manuscript's final draft. All authors discussed the results and contributed to the final text of the manuscript.

Conflict of Interest

The authors declare no conflict of interest.

Acknowledgements

This work/article is an output of a research project (HSE-BR-2025-010) implemented as part of the Basic Research Program at HSE University.

References

Ackerman, R., & Lauterman, T. (2012). Taking reading comprehension exams on screen or on paper? A metacognitive analysis of learning texts under time pressure. *Computers in Human Behavior*, *28*(5), 1816–1828. https://doi.org/10.1016/j.chb.2012.04.023

Aiken, L.S., & West, S.G. (1991). *Multiple regression: Testing and interpreting interactions.* Sage Publications.

Akpen, C.N., Asaolu, S., Atobatele, S., Okagbue, H., & Sampson, S. (2024). Impact of online learning on student's performance and engagement: A systematic review. *Discover Education, 3*(1), 205. https://doi.org/10.1007/s44217-024-00253-0

Austin, P.C., & Steyerberg, E.W. (2015). The number of subjects per variable required in linear regression analyses. *Journal of Clinical Epidemiology, 68*(6), 627–636. https://doi.org/10.1016/j.jclinepi.2014.12.014

Azevedo, R., & Hadwin, A.F. (2005). Scaffolding self-regulated learning and metacognition: Implications for the design of computer-based learning environments. *Instructional Science, 33*(5–6), 367–379. https://doi.org/10.1007/s11251-005-1272-9

Bannert, M., & Mengelkamp, C. (2013). Scaffolding hypermedia learning through metacognitive prompts. In R. Azevedo & V. Aleven (Eds.), *International handbook of metacognition and learning technologies* (pp. 171–186). Springer. https://doi.org/10.1007/978-1-4419-5546-3_12

Barak, M., Hussein-Farraj, R., Dori, Y. (2016). On-campus or online: Examining self-regulation and cognitive transfer skills in different learning settings. *International Journal of Educational Technology in Higher Education, 13,* 35. https://doi.org/10.1186/s41239-016-0035-9

Basar, Z.M., Mansor, A.N., Jamaludin, K.A., & Alias, B.S. (2021). The effectiveness and challenges of online learning for secondary school students – a case study. *Asian Journal of University Education, 17*(3), 119. https://doi.org/10.24191/ajue.v17i3.14514

Beege, M., & Ploetzner, R. (2025). Learning from interactive video: The influence of self-explanations, navigation, and cognitive load. *Instructional Science, 53*(1), 99–119. https://doi.org/10.1007/s11251-024-09693-5

Benjamini, Y., & Hochberg, Y. (1995). Controlling the false discovery rate: A practical and powerful approach to multiple testing. *Journal of the Royal Statistical Society* Series B (Statistical Methodology), *57*(1), 289–300. https://doi.org/10.1111/j.2517-6161.1995.tb02031.x

Berthold, K., Nückles, M., & Renkl, A. (2007). Do learning protocols support learning strategies and outcomes? The role of cognitive and metacognitive prompts. *Learning and Instruction, 17*(5), 564–577. https://doi.org/10.1016/j.learninstruc.2007.09.007

Brown, G.T.L., & Abdulnabi, H.H.A. (2017). Evaluating the quality of higher education instructor-constructed multiple-choice tests: Impact on student grades. *Frontiers in Education, 2*, 24. https://doi.org/10.3389/feduc.2017.00024

Costley, J., Kapuza, A., Gorbunova, A., & Shcheglova, I. (2025). How adding structure reduces complexity: More interconnections in concept maps do not increase cognitive load. *Instructional Science, 53*(5), 1243–1262. https://doi.org/10.1007/s11251-025-09736-5

Credé, M., & Phillips, L.A. (2011). A meta-analytic review of the Motivated Strategies for Learning Questionnaire. *Learning and Individual Differences, 21*(4), 337–346. https://doi.org/10.1016/j.lindif.2011.03.002

de Bruin, A.B.H., Roelle, J., Carpenter, S.K., Baars, M., & Efg-Mre (2020). Synthesizing Cognitive Load and Self-Regulation Theory: A theoretical framework and research agenda. *Educational Psychology Review, 32*(4), 903–915. https://doi.org/10.1007/s10648-020-09576-4

Dignath, C., Van Ewijk, R., Perels, F., & Fabriz, S. (2023). Let learners monitor the learning content and their learning behavior! A meta-analysis on the effectiveness of tools to foster monitoring. *Educational Psychology Review, 35*(2), 62. https://doi.org/10.1007/s10648-023-09718-4

Dörrenbächer-Ulrich, L., Weißenfels, M., Russer, L., & Perels, F. (2021). Multimethod assessment of self-regulated learning in college students: Different methods for different components? *Instructional Science*, *49*(1), 137–163. https://doi.org/10.1007/s11251-020-09533-2

Efron, B., & Tibshirani, R. (1994). *An introduction to the Bootstrap*. Chapman & Hall/CRC. https://doi.org/10.1201/9780429246593

Engelmann, K., Bannert, M., & Melzner, N. (2021). Do self-created metacognitive prompts promote short- and long-term effects in computer-based learning environments? *Research and Practice in Technology Enhanced Learning*, *1,* 16. https://doi.org/10.1186/s41039-021-00148-w

Galperin, P.Ya. (1969). Stages in the development of mental acts. In M. Cole & I. Maltzman (Eds.) *A handbook of contemporary Soviet psychology* (pp. 249-273). Basic Books.

Glogger-Frey, I., Gaus, K., & Renkl, A. (2016). Learning from direct instruction: Best prepared by several self-regulated or guided invention activities? *Learning and Instruction*, *51,* 26–35. https://doi.org/10.1016/j.learninstruc.2016.11.002

Gorbunova, A., Adamovich, K.A., Savelyev, A., & Costley, J. (2026). Behavior patterns characterize students' choices and relate to cognitive load and performance in learner-controlled environments. *The Internet and Higher Education*, *69*, 101073. https://doi.org/10.1016/j.iheduc.2025.101073

Greenberg, K., & Zheng, R. (2023). Revisiting the debate on germane cognitive load versus germane resources. *Journal of Cognitive Psychology*, *35*(3), 295–314. https://doi.org/10.1080/20445911.2022.2159416

Guo, L. (2022). The effects of self-monitoring on strategy use and academic performance: A meta-analysis. *International Journal of Educational Research*, *112,* 101939. https://doi.org/10.1016/j.ijer.2022.101939

Harrell, F.E. (2015). *Regression modeling strategies: With applications to linear models, logistic and ordinal regression, and survival analysis* (2nd ed.). Springer. https://doi.org/10.1007/978-3-319-19425-7

Huff, J.D., & Nietfeld, J.L. (2009). Using strategy instruction and confidence judgments to improve metacognitive monitoring. *Metacognition and Learning*, *4*(2), 161–176. https://doi.org/10.1007/s11409-009-9042-8

Jung, J., Shin, Y., Chung, H., & Fanguy, M. (2024). The effects of pre-training types on cognitive load, self-efficacy, and problem-solving in computer programming. *Journal of Computing in Higher Education*, *37*(3), 809–830. https://doi.org/10.1007/s12528-024-09407-3

Kalyuga, S., & Singh, A. (2015). Rethinking the boundaries of cognitive load theory in complex learning. *Educational Psychology Review*, *28*(4), 831–852. https://doi.org/10.1007/s10648-015-9352-0

Karpov, A.A., Karpov, A.V., Karabushchenko, N.B., & Ivashchenko, A.V. (2017). The interconnection of learning ability and the organization of metacognitive processes and traits of personality. *Psychology in Russia: State of the Art*, *10*(1), 67–79. https://doi.org/10.11621/pir.2017.0105

Krieglstein, F., Beege, M., Rey, G.D., Sanchez-Stockhammer, C., & Schneider, S. (2023). Development and validation of a theory-based questionnaire to measure different types of cognitive load. *Educational Psychology Review*, *35*(1), 9. https://doi.org/10.1007/s10648-023-09738-0

Löffler, E., Von Der Linden, N., & Schneider, W. (2016). Influence of domain knowledge on monitoring performance across the life span. *Journal of Cognition and Development*, *17*(5), 765–785. https://doi.org/10.1080/15248372.2016.1208204

Merkt, M., & Bodemer, D. (2024). Learning with videos: Do task instructions and the availability of a pause button matter? *Journal of Computer Assisted Learning*, *40*(6), 2856–2871. https://doi.org/10.1111/jcal.13044

Mou, T.-Y. (2023). Online learning in the time of the COVID-19 crisis: Implications for the self-regulated learning of university design students. *Active Learning in Higher Education*, *24*(2), 185-205. https://doi.org/10.1177/14697874211051226

Müller, N.M., & Seufert, T. (2018). Effects of self-regulation prompts in hypermedia learning on learning performance and self-efficacy. *Learning and Instruction*, *58*, 1–11. https://doi.org/10.1016/j.learninstruc.2018.04.011

Nakagawa, S. (2004). A farewell to Bonferroni: The problems of low statistical power and publication bias. *Behavioral Ecology*, *15*(6), 1044–1045. https://doi.org/10.1093/beheco/arh107

Nelson, T.O., & Narens, L. (1990). Metamemory: A theoretical framework and new findings. *The Psychology of Learning and Motivation*, *26*, 125–141. https://doi.org/10.1016/S0079-7421(08)60053-5

Perikova, E. I., & Byzova, V. M. (2022). Faktornaia struktura russkoiazychnoi versii oprosnika "Metakognitivnaia vkliuchennost v deiatelnost" [Factor structure of the Russian version of the "Metacognitive Awareness Inventory"]. *Kulturno-istoricheskaia psikhologiia* [Cultural-Historical Psychology], *18*(2), 116–126. https://doi.org/10.17759/chp.2022180213

Pikhart, M., Klimova, B., Cierniak-Emerych, A., & Dziuba, S. (2023). A comparative analysis of perceived advantages and disadvantages of online learning. *Behavioral Sciences*, *13*(3), 262. https://doi.org/10.3390/bs13030262

Pintrich, P.R. (2000). The role of goal orientation in self-regulated learning. In M. Boekaerts, P. R. Pintrich, & M. Zeidner (Eds.), *Handbook of self-regulation* (pp. 451–502). Academic Press. https://doi.org/10.1016/B978-012109890-2/50043-3

Pham, N.S. (2022). The effectiveness of teaching and learning online: A study on HUFI English-majored students. *International Journal of TESOL & Education*, *2*(3), 1–12. https://doi.org/10.54855/ijte.22231

Podolsky, A.I. (2023). Vstupitelnoe slovo [Foreword]. In P. Ya. Galperin, *Psikhologiia: predmet i metod. Izbrannye psikhologicheskie trudy* [Psychology: Subject and method. Selected psychological works] (pp. 13–32). Moscow University Press.

Rogozin, D., & Ipatova, A. (2020). Metodicheskii audit monitoringovykh sotsiologicheskikh issledovanii [Methodological audit of monitoring social research]. *SSRN Electronic Journal*. https://doi.org/10.2139/ssrn.3663255

Schraw, G., Crippen, K.J., & Hartley, K. (2006). Promoting self-regulation in science education: Metacognition as part of a broader perspective on learning. *Research in Science Education*, *36*(1–2), 111–139. https://doi.org/10.1007/s11165-005-3917-8

Schraw, G., & Dennison, R.S. (1994). Assessing metacognitive awareness. *Contemporary Educational Psychology*, *19*(4), 460–475. https://doi.org/10.1006/ceps.1994.1033

Schunk, D.H. (1982). Progress self-monitoring. *The Journal of Experimental Education*, *51*(2), 89–93. https://doi.org/10.1080/00220973.1982.11011845

Seufert, T. (2018). The interplay between self-regulation in learning and cognitive load. *Educational Research Review*, *24*, 116–129. https://doi.org/10.1016/j.edurev.2018.03.004

Shi, Y., & Qu, S. (2022). Analysis of the effect of cognitive ability on academic achievement: Moderating role of self-monitoring. *Frontiers in Psychology*, *13*, 996504. https://doi.org/10.3389/fpsyg.2022.996504

Sidi, Y., & Ackerman, R. (2024). Opting out as an untapped resource in instructional design: Review and implications. *Educational Psychology Review*, *36*(2), 41. https://doi.org/10.1007/s10648-024-09879-w

Singh, S., Loh, J. Y., & Lin, P.K.F. (2024). Impact of online learning experience during COVID-19: Exploring the relationship between academic resilience, academic self-efficacy, and perceived social support. In M. Zangeneh (Ed.), *Essentials in health and mental health: Unlocking the keys to wellness* (pp. 39–56). Springer Nature. https://doi.org/10.1007/978-3-031-56192-4_3

Sistek-Chandler C.M. (2019). Exploring online learning through synchronous and asynchronous instructional methods. In *Advances in mobile and distance learning book series*. IGI Global. https://doi.org/10.4018/978-1-7998-1622-5

Skulmowski, A., & Xu, K.M. (2021). Understanding cognitive load in digital and online learning: A new perspective on extraneous cognitive load. *Educational Psychology Review*, *34*(1), 171–196. https://doi.org/10.1007/s10648-021-09624-7

Sweller, J., van Merriënboer, J.J., & Paas, F.G. (1998). Cognitive architecture and instructional design. *Educational Psychology Review*, *10*, 251–296. https://doi.org/10.1023/A:1022193728205

Sweller, J., Van Merriënboer, J.J., & Paas, F. (2019). Cognitive architecture and instructional design: 20 years later. *Educational Psychology Review*, *31*(2), 261–292. https://doi.org/10.1007/s10648-019-09465-5

Tabachnick, B.G., & Fidell, L.S. (2006). *Using multivariate statistics* (5th ed.). Pearson.

Talyzina, N.F. (1981) *The psychology of learning*. Progress.

Viberg, O., Khalil, M., & Baars, M. (2020). Self-regulated learning and learning analytics in online learning environments: A review of empirical research. In *Proceedings of the Tenth International Conference on Learning Analytics & Knowledge (LAK '20)*, 524–533. https://doi.org/10.1145/3375462.3375483

Winne, P.H., & Hadwin, A.F. (1998). Studying as self-regulated learning. In D.J. Hacker, J. Dunlosky, & A.C. Graesser (Eds.), *Metacognition in educational theory and practice* (pp. 277–304). Lawrence Erlbaum Associates.

Wong, J., Baars, M., Davis, D., Van Der Zee, T., Houben, G.J., & Paas, F. (2019). Supporting self-regulated learning in online learning environments and MOOCs: A systematic review. *International Journal of Human–Computer Interaction*, *35*(4-5), 356–373. https://doi.org/10.1080/10447318.2018.1543084

Zimmerman, B.J. (2000). Attaining self-regulation: A social cognitive perspective. In M. Boekaerts, P.R. Pintrich, & M. Zeidner (Eds.), *Handbook of self-regulation* (pp. 13–39). Academic Press.

Zimmerman, B.J., & Moylan, A.R. (2009). Self-regulation: Where metacognition and motivation intersect. In D. J. Hacker, J. Dunlosky, & A. C. Graesser (Eds.), *Handbook of metacognition in education* (pp. 299–315). Routledge.

Zimmerman, B.J., & Schunk, D.H. (2011). Self-regulated learning and performance: An introduction and an overview. In B. J. Zimmerman & D. H. Schunk (Eds.), *Handbook of self-regulation of learning and performance* (pp. 1–12). Routledge.

Original manuscript received May 5, 2026
Revised manuscript accepted August 27, 2026
First published online September 1, 2026

To cite this article: Koval, A.A., Gorbunova, A.Y., Podolsky, A.I., Getman, A.V. (2026). Self-Monitoring Skills in Online Learning and the Role of Metacognitive Prompts: An Exploratory Study. *Psychology in Russia: State of the Art*, *19*(3), 89–105. DOI: 10.11621/pir.2026.0305

Psychology in Russia: State of the Art
Volume 19, Issue 3, 2026

Russian Psychological Society
Lomonosov Moscow State University

PSYCHOMETRICS

Psychometric Characteristics of the Russian-Language Version of the Posttraumatic Growth and Depreciation Inventory — Expanded Version (PTGDI-X)

Alexandra M. Shishkova[a*], Victor V. Bocharov[a,b,c], Tatiana V. Valieva[d], Richard G. Tedeschi[e]

[a] *V.M. Bekhterev National Research Medical Center for Psychiatry and Neurology, Saint-Petersburg, Russia*
[b] *Saint Petersburg State Pediatric Medical University, Russia*
[c] *Saint Petersburg State University, Russia*
[d] *Ural Federal University named after the first President of Russia B.N. Yeltsin, Ekaterinburg, Russia*
[e] *Boulder Crest Institute for Posttraumatic Growth, Virginia, USA*

* Corresponding author. E-mail: shishaspb@mail.ru

Background. The issue of caring for chronically and seriously ill patients is becoming increasingly relevant all over the world. Relatives caring for such patients typically perform a significant number of functions to maintain the well-being of their ailing loved ones and are often exposed to significant psychological distress or even trauma during the caregiving process. Various tools have been developed to assess such distress, family burden, and burnout among relatives of chronically ill patients. At the same time, there is an acute shortage of tools for assessing the constructive aspects of the transformation of an individual's values, beliefs and meanings under conditions of chronic stress. One such tool, which allows for the simultaneous assessment of both positive and negative changes in mental functioning resulting from exposure to traumatic events, is the "Posttraumatic Growth and Posttraumatic Depreciation Inventory — Expanded version" (PTGDI-X).

Objective. To define the factor structure and psychometric characteristics of the Russian-language version of the PTGDI-X.

Design. The sample consisted of 404 Russian-speaking informal caregivers (*M* age = 48.95 years, *SD* = 13.3; 89 males and 315 females). A quantitative approach with non-experimental, cross-sectional design and instrumental scope.

Keywords: posttraumatic growth, posttraumatic depreciation, informal caregivers, factor structure, construct validity, assessment

© Shishkova, A.M., Bocharov, V.V., Valieva, T.V., Tedeschi, R.G., 2026

ISSN 2074-6857 (Print) / ISSN 2307-2202 (Online)
http://psychologyinrussia.com

Results. Analysis of the factor structure using Confirmatory Factor Analysis (CFA) indicates that the models of Posttraumatic Growth (PTG) and Posttraumatic Depreciation (PTD) fit the empirical data and confirms the construct validity of the analyzed version of the questionnaire. A significant inverted U-shaped quadratic relationship was found between the scores on the overall PTG scale and the overall PTD scale. The Russian-language version of the questionnaire demonstrated good psychometric characteristics in terms of reliability (Cronbach's alpha coefficients were .956 for the PTG-total and .942 for the PTD-total scales; McDonald's omega values were .964 and .952, respectively), internal construct validity, and convergent and divergent cross-validity.

Conclusion. The Russian-language version of the PTGDI-X can be used to study posttraumatic growth and depreciation arising in individuals in various stressful and psychologically traumatic situations.

Introduction

In recent decades, the problems and experiences of relatives caring for the chronically ill have increasingly attracted the attention of the scientific community. In particular, this is related to the growing population of individuals suffering from chronic non-communicable diseases and in need of constant care (Ansah & Chiu, 2023; Antonova et al., 2021; Hacker, 2024).

The term "informal caregiver" (IC) is most commonly used in the literature to refer to those who are involved in taking care of chronically or seriously ill persons. Based on a publications analysis, we developed the following definition: "an IC is someone who has a social relationship (such as a spouse, parent, child, another relative, neighbor, friend, etc.) with the care recipient (a dependent person in need of care due to various circumstances) and provides unpaid care on a permanent basis or when a necessity arises" (Shishkova & Bocharov, 2021, p. 19). ICs of individuals with severe or chronic illnesses make a substantial contribution to the development of non-hospital forms of care for this patient population; however, their role within the healthcare system is often underestimated (Barello et al., 2019; Shishkova & Bocharov, 2021)

It is crucial to highlight the lack of a system for psychological assessment for ICs of chronically ill patients within the domestic healthcare model, particularly within the framework of specialized public institutions. At the same time, evaluating the psychological well-being of family members, providing care for individuals with severe or chronic somatic and/or mental disorders, appears to be critically important, given the considerable amount of empirical data gathered to date demonstrating that such care is often associated with exposure to acute and/or long lasting (chronic) psychological trauma and stress (Cejalvo et al., 2025; Choi et al., 2024; Shishkova & Bocharov, 2021).

Psychological trauma is an experience that feels very painful or distressing, leading to short- or long-term mental and physical limitations or an inability to function (Feriante & Sharma, 2023). Traumatic events usually involve threatened death or severe injury, either to the individuals themselves or to a loved one (Cureton, 2014).

Acute trauma arises from a single traumatic incident, whereas chronic trauma develops from repeated or prolonged exposure to ongoing traumatic experiences (Feriante & Sharma, 2023). The diagnosis of a life-threatening illness in a close relative may itself constitute a psychotraumatic event for the patient's family members, potentially resulting in serious emotional distress and difficulties in mental functioning (Choudry et al., 2024). Empirical evidence suggests that family members and close associates of individuals with impaired health conditions (*e.g.*, those affected by COVID-19) may experience higher levels of posttraumatic stress than the patients themselves (Moreta-Herrera et al., 2025). Exploring the mechanisms and moderators underlying this phenomenon could provide important insights into theoretical models of trauma transmission and clinical interventions aimed at supporting affected families.

Financial burdens, disruption of or forced curtailment of hobbies and social contacts, feelings of loss concerning premorbid relationships with a person close to one, as well as concerns about his/her health, safety, and future are often experienced by ICs of people with chronic illnesses. Negative consequences of exposure to caregivers' stress encompass various emotional disorders, deteriorating health and well-being (Alves et al., 2019; Barello et al., 2019). Depressive disorders, severe anxiety and sleep disturbances are frequently seen in ICs (Alves et al., 2019; Choi et al., 2024; Shishkova & Bocharov, 2021); even posttraumatic stress reactions are possible (Lim et al., 2024). Additionally, studies include negative consequences of the care of chronically ill patients that can be described as burnout, and may appear as the form of a feeling of chronic fatigue and inability to continue care, a distant or cynical attitude towards the patient, and a sense of worthlessness regarding one's efforts to overcome the disease (Ci Soh et al., 2025; Shishkova & Bocharov, 2021).

Thus, the stress-oriented approach focuses primarily on the negative consequences of care. This may naturally, to some extent, determine the direction of development of the psycho-diagnostic tools used in the assessment of the psychological consequences experienced by ICs. For example, various tools have been developed to assess distress, family burden and burnout of ICs of chronically ill patients (Shishkova et al., 2021; Tamarana et al., 2023; Zhong et al., 2024).

At the same time, researchers have focused mainly on the negative effects of traumatic events, while the possible positive outcomes of traumatic experiences were largely overlooked and their study remained has limited for many years, partly due to the lack of suitable psychodiagnostic tools.

The Model of Posttraumatic Growth and Posttraumatic Depreciation

In connection with the changes of scientific paradigm toward the framework of positive psychology, one of the most promising areas of research recently is the study not only of the negative consequences, but also of positive outcomes, which can arise in an individual in stressful or psychotraumatic situations. ICs may perceive a loved one's illness as an opportunity to be stronger and acquire new skills or as an event that gives a sense of purpose in life and spiritual growth, etc. (Palacio & Limonero, 2020; Sergienko, 2019).

The functional-descriptive model of posttraumatic growth (PTG), proposed by Tedeschi and Calhoun (2004), seems to be the crucial theoretical framework that has made it possible to conceptualize the observed phenomena. According to this model, posttraumatic growth (PTG) is the product of an individual's efforts under conditions in which adaptation to the situation may seem impossible, and where the unavoidability of change necessitates the transformation of fundamental worldviews, beliefs, and personal goals. The construct has so far been confirmed in empirical studies on different populations, including ICs of patients with oncological illnesses, children with disabilities, and others (Palacio & Limonero, 2020; Sergienko, 2019; Taku et al., 2021).

The Posttraumatic Growth Inventory (PTGI) (Tedeschi & Calhoun, 1996), developed by those authors within the functional-descriptive model, was for a long time the most widely used tool for assessing positive changes appears in an individual's values, beliefs, and meanings arising from ordeals. The PTGI consists of 21 statements and has a five-factor structure that has been confirmed in different cultural groups, showing only minor variations (Exenberger et al., 2019; Weiss & Berger, 2010). The PTGI includes five dimensions, which can be defined as follows: Personal Strength — an increased sense of resilience and self-reliance; New Possibilities — greater openness to new opportunities and life paths; Relating to Others — deeper compassion and closer relationships; Appreciation of Life — increased gratitude and re-evaluation of priorities; and Spiritual Change — greater spiritual or existential awareness. These factors constitute the corresponding scales of the questionnaire and represent different areas of personal growth.

Further development of the concept suggested by Tedeschi and Calhoun is evident in the formation of the concept of posttraumatic depreciation, which reflects negative changes resulting from an individual's exposure to stressful or traumatic situations. These changes, according to the authors, take place in the same areas of personal functioning mentioned above (Baker et al, 2008; Cann et al., 2010).

The expansion of the theoretical construct through the notion of posttraumatic depreciation necessitated corresponding transformations of the measurement tool. For this reason, the PTGI was supplemented with negatively formulated statements and renamed the PTGI-42: Growth and Depreciation (Baker et al, 2008). This was developed in response to criticism that the PTGI was biased toward agreement with positive experiences. Studies have indicated that PTG is associated with positive experiences, whereas PTD is linked to negative ones (Barrington & Shakespeare-Finch, 2013). The emerging possibility of simultaneously examining multidirectional transformations occurring in the personal axiological field under the influence of a psychological trauma has significantly expanded the diagnostic potential of the assessment tool while simultaneously providing an opportunity for a deeper understanding of the mechanisms involved in response to stress — in particular, the relationship between negative and positive personal transformations. The emergence of these new possibilities has generated considerable interest among researchers, as reflected in numerous studies aimed at investigation of the interrelation of the constructs between PTG and PTD (Cann et al., 2010; Barrington & Shakespeare-Finch, 2013; Kunz et al., 2019). Thus, the comprehensive understanding of the mechanisms of personal responses to psychologically traumatic situations is impossible without

providing respondents with the opportunity to simultaneously assess both positive and negative changes resulting from stress.

As it currently stands, Tedeschi and co-authors have proposed an expanded version of the inventory: Posttraumatic Growth and Posttraumatic Depreciation Inventory - Expanded version PTGDI-X (Taku et al., 2020). The PTGDI-X is based on the PTGI-X, which was developed to improve the assessment of Spiritual and Existential Change (Tedeschi, et al., 2017). The PTGDI-X was then used in ten countries of varying cultural backgrounds in order to determine the broad applicability of the PTG and PTD concepts. The factor structure and conceptual relationships of the PTGDI-X has been examined in several countries in addition to the ten-country sample on which it was originally based. These countries include Japan (Oshiro, et al., 2019), Italy (Romeo, et al., 2023), Germany (Platte et al., 2023), and Sierra Leone (Steidl, et al., 2024). The five-factor structure failed to replicate in Sierra Leone as it did in other countries, leaving the open the possibility that varying cultures may conceptualize growth and depreciation in different ways. Therefore, when used with a Russian population, it is useful to determine whether the proposed factor structure applies, especially since Russia was not one of the ten countries on which the measure was based. In addition to testing the factor structure, the relationship between PTG and PTD is an important theoretical and clinical issue to address. PTG and PTD appear to be separate dimensions of post-trauma experience (Manis, et al., 2025; Taku, et al., 2021), rather than opposing poles of a unidimensional construct. Again, cultural factors could potentially be involved in this relationship.

The Present Study

For Russian-speaking respondents, only an older version of the PTGI (Magomed-Eminov, 2008) is currently available, which lacks the additional items involved in the PTGI-X. Therefore, the authors of this article conducted a cross-cultural adaptation and primary approbation of the PTGDI-X on a sample of ICs of patients with cerebral pathology (Shishkova et al., 2024).

The aim of this work was to verify the factor structure and assess the psychometric characteristics of the Russian-language version of the PTGDI-X on a sample of ICs of patients with severe and/or chronic mental, neurological, and somatic diseases. Differences in the clinical presentation and progression of the patient's illness impact the type and severity of psychotraumatic and stressful impacts experienced by family caregivers. In other words, caregiving can affect several aspects of life and relationships, including anticipating the loss of a loved one, changes in social status, and the necessity to adjust life goals, all of which can lead to significant stress. Consequently, this sample is appropriate for studying the effects of exposure to stressful and traumatic events. In our opinion, ICs can be considered one of the most important models for studying posttraumatic and stress-related responses due to their exposure to both long-term and acute psychological stress. Therefore, using the PTGDI-X model with informal caregivers will also help assess whether the instrument is applicable not only to those caring for chronically ill persons, but also to other groups exposed to psychological trauma or ongoing stress.

Methods

The study applied a quantitative approach with a non-experimental, cross-sectional design and an instrumental scope (Slater & Hasson, 2025).

Participants

The study included 404 ICs of patients suffering from various mental, neurological, and somatic disorders. Among those examined were 55 ICs of patients with schizophrenia spectrum disorders (F20, F21, F25), 45 ICs of patients with affective disorders (F31, F32, F34), 53 with anxiety, depressive and somatoform neurotic disorders (F41.0, F41.1, F41.2, F45.0), 45 with epilepsy (G40.0 - 40.4), 13 with multiple sclerosis (G35.0), 40 with Parkinsonism (G20), 52 ICs of patients with consequences of acute cerebrovascular event (I69), 52 ICs of patients with primary/secondary malignant or benign neoplasms of the brain and/or spinal cord (C71.1- C71.8; D32.1; D33.4), 49 ICs of patients with head and spinal cord injuries (S06; T91.3; T90.5). Patients were diagnosed based on the ICD-10.

The requirement for selecting these respondent groups was their involvement in taking care of patients with severe and/or chronic disease (psychiatric, neurological, or physical). ICs had to be over 18 years, able to provide informed consent, and willing to follow the study procedures. Participants were excluded if they had a medical condition severe enough to prevent them from providing informed consent or understanding the study procedure.

Sampling of respondents was based on the requirement of achieving representation of families having members who were living in circumstances related to different care models. In the present article, a care model is considered as the combination of psychological, physical, material, and other forms of support provided by a relative to an ill loved one in order to maintain the latter's well−being. The extent of help provided is dependent on the level of dysfunction exhibited by the patient, both physically and psychologically. Patient care could be characterized by attributes like the level of somato-neurological dysfunction and the inability of the patient to interact with the caregiver. Thus, in the case of pronounced somatic or neurological impairments, caregiver stress was associated more with caregiving tasks such as hygienic procedures. When dealing with psychiatric disorders, caregiver stress can be associated with stigma and self stigma, communication problems, and family conflict. In our current sample, there were caregivers who dealt with the threat of death of their relatives, while others dealt with challenging behaviors and associated conflicts. At the same time, all caregivers can be assigned to one category, since their patients' diseases entailed a great danger of disability and caused stress among family members who gave them care. ICs might encounter many challenges when dealing with various kinds of diseases and even when dealing with the same disorder, informal caregivers can go through heterogeneous challenges and emotional reactions, including the fear of losing someone close to them, which is possible not only in oncology but also in multiple sclerosis and depression.

The inclusion of ICs of patients with depressive and anxiety neurotic disorders, despite their psychogenic and functional nature, was determined by the severity of

the clinical picture in patients, characterized by pronounced disturbances in the emotional-volitional sphere, difficulties in daily life and self-care, and the presence of somatic correlates of anxiety and depression.

A comparative analysis of respondent groups based on the nosology of the care recipient was not an aim of this study and will be examined by us in future publications. The process of validation of the instrument was done without segregating the subjects into groups. The core concept of the article is to validate, through the use of a caregiver model, the PTGDI-X as an instrument for studying PTG and PTD.

Table 1

Demographic and Clinical Characteristics of the Sample

Sociodemographic and clinical characteristics	***n* (%):**
Relatives' gender	
Male	89 (22%)
Female	315 (78%)
Relatives' age / mean (M ± SD)	48.95 ± 13.3
Relationship to patient	
Father	19 (4.6%)
Mother	176 (43.5%)
Husband	52 (12.7%)
Wife	81 (20.5%)
Son	15 (3.7%)
Daughter	51 (12.6%)
Sister	7 (1.7%)
Brother	3 (.7%)
Employment status	
Not working	115 (28.5%)
Working	289 (71.5%)
Patients' age / mean (*M* ± *SD*)	40.57 ± 18.8
Patients' gender	
Male	204 (50.5%)
Female	200 (49.5%)
Age of the patients' disease onset / mean (*M* ± *SD*)	33.82 ± 20.3
Duration of the disease / mean (*M* ± *SD*)	6.68 ± 7.9
Living with patient	
Together	274 (67.8%)
Separate	130 (32.2%)

The sociodemographic and clinical characteristics of ICs and patients care by them are shown in *Table 1*.

The mean age of the ICs was 48.95 (*SD* = 13.3) years. Most of the ICs were women — 78% (*n* = 315). By relation, the respondents were predominantly mothers — 43.5% (*n* = 176), fathers were 4.6% (*n* = 19), wives — 20.5% (*n* = 81), husbands — 12.7% (*n* = 52), daughters — 12.6% (*n* = 51), sons — 3.7% (*n* = 15), sisters — 1.7% (*n* = 7), brothers — .7% (*n* = 3). Most of the surveyed relatives had higher education — 69.8% (*n* = 282), and were employed — 71.5% (*n* = 289). At the time of conducting psychodiagnostics most of the relatives lived with the cared-for patients — 67.8% (*n* = 274). The median age of the cared-for patients was 40.57 (*SD* = 18.8) years. Among them 50.5% (*n* = 204) were men and 49.5% (*n* = 200) women. The average age of onset of the patient's illness was 33.82 (*SD* = 20.3) years, the duration of the illness was 6.68 (*SD* = 7.9) years.

Procedure

The cross-cultural adaptation of the PTGDI-X inventory was conducted according to all the necessary stages of the procedure (Beaton et al., 2000; Epstein et al., 2015) and was described in a previous publication (Shishkova et al., 2024). The patients cared for by ICs were treated in V.M. Bekhterev National Medical Research Centre for Psychiatry and Neurology and Saint Petersburg clinical scientific and practical center for specialized types of medical care (oncological). All participants provided voluntary informed consent prior to the study. During a face-to-face meeting, a psychologist explained the study procedures and provided relevant information to each relative before obtaining their informed consent. Participants did not receive any payment for participating and were assured that their responses would be kept confidential and that all data would be anonymized prior to analysis. Recruitment and data collection took place between April 2023 and November 2025.

The diagnostic procedure for ICs consisted of: (1) a semi-structured interview during which we sought to establish the most trusting atmosphere possible while discussing the respondents' current life situation and collecting socio-demographic and clinical characteristics of the ICs and their care recipients, (2) psychological assessment using quantitative methods, and (3) a "feedback" procedure.

Questionnaires

The original *PTGDI-X* is a 50-item self-report questionnaire designed to assess both PTG and PTD dimensions (Taku et al., 2020). Each item is evaluated on a 6-point Likert-type scale ranging from 0 to 5, where higher scores reflect greater levels of either growth or depreciation. Beyond the composite scores for each dimension, the instrument allows for the calculation of scores across five corresponding subscales for PTG and PTD: Relating to Others, New Possibilities, Personal Strength, Spiritual and Existential Change, and Appreciation of Life.

To study convergent and discriminant validity, the following instruments were used: *The World Assumption Scale (WAS)* by Janoff-Bulman adapted by Padun and

Kotelnikova (2008); the *Symptom Checklist SCL-90-Revise*d adapted by Tarabrina (2001), and the *Level of Relatives' Emotional Burnout Questionnaire (LREB)* (Shishkova et al., 2021). *The World Assumption Scale (WAS)*, originally developed by Janoff-Bulman (1989), measures core beliefs regarding the benevolence of the world, the meaning of events, and self-worth. The Russian adaptation by Padun and Kotelnikova retains the original multidimensional structure and assesses five key assumptions: Benevolence of the World, Justice, Self-Worth, Luck, and Perceived Control. The adapted version shows good reliability, with Cronbach's α coefficients ranging from .70 to .85 across the subscales.

The Symptom Checklist-90-Revised (SCL-90-R), originally developed by Derogatis (1973), is a self-report measure designed to assesses current psychological symptoms in clinical and non-clinical populations. The Russian adaptation, validated by Tarabrina (2001), retains the original 90-item structure and evaluates nine symptom dimensions: Somatization, Obsessive-Compulsive, Interpersonal Sensitivity, Depression, Anxiety, Hostility, Phobic Anxiety, Paranoid Ideation, and Psychoticism, along with three global distress indices (GSI, PSDI, PST). The adapted version demonstrates satisfactory psychometric properties, with Cronbach's α coefficients ranging from .73 to .88.

The Level of Relatives' Emotional Burnout Questionnaire (LREB), developed by Shishkova et al. (2021), is a self-report measure assessing emotional burnout among relatives caring for chronically ill patients. In addition to the core burnout components—exhaustion, depersonalization, and reduced personal accomplishment—the instrument assesses personal resources related to caregiving, including dedication and self-efficacy. The questionnaire comprises eight scales: Vigor, Dedication, Resource, Self-Efficacy, Exhaustion, Depersonalization, Destruction, and Inefficacy. The subscales showed satisfactory internal consistency, with Cronbach's α coefficients ranging from .70 to .86.

Data Analysis

The statistical analysis was conducted in several sequential phases. First, confirmatory factor analysis (CFA) was performed to assess the fit of the proposed factor structure to the observed data using an appropriately justified estimation method. Second, internal consistency was assessed using Cronbach's alpha and McDonald's omega and internal construct validity was evaluated using Spearman's rank correlations among the PTGDI-X subscales. The relationship between the overall PTG and PTD scores was examined using nonlinear regression analysis. In the final phase, external construct validity was assessed by examining Spearman's correlations between the PTGDI-X scales and the WAS, SCL-90-R, and LREB scales. Descriptive statistical methods were calculated. All analyses were conducted using the R statistical programming environment, version 4.6.0 (R Core Team, 2026). Confirmatory factor analysis was performed using the *lavaan* package, version .6-12 (Rosseel, 2012). The PTG and PTD factor structures were visualized using yEd Graph Editor, version 3.23.2 (yWorks GmbH, 2023).

Results

Descriptive Statistics

The descriptive statistics are presented in *Table 2*.

The revealed positive skewness across all scales, negative kurtosis for the PTG scales and predominantly positive kurtosis for the PTD scales, as well as large values of standard deviation, necessitate the application of robust methods in subsequent data processing.

Table 2

Descriptive Statistics

Scales	Descriptive Statistics				
	Mean	**Median**	***SD***	**Skewness**	**Kurtosis**
PTG/Relating to Others	12.96	12.0	9.19	.289	–.887
PTG/New Possibilities	7.72	6.0	6.47	.707	–.511
PTG/Personal Strength	7.31	7.0	5.70	.357	–.988
PTG/Spiritual and Existential Change	8.75	8.0	7.88	.671	–.484
PTG/Appreciation of Life	6.65	6.0	4.60	.121	–1.139
PTG/Total	43.40	42.0	30.51	.409	–.727
PTD/Relating to Others	5.79	4.0	6.34	1.205	.933
PTD/New Possibilities	5.86	4.5	5.83	.733	–.564
PTD/Personal Strength	3.78	2.0	4.37	1.233	.963
PTD/Spiritual and Existential Change	4.83	3.0	5.80	1.351	1.253
PTD/Appreciation of Life	2.25	1.0	3.03	1.538	1.990
PTD/Total	22.52	15.0	22.44	1.062	.458

Note. Standard error of skewness = .121; standard error of kurtosis = .242; SD - standard deviation

Factor Structure

The factor structure of the inventory was examined using confirmatory factor analysis (CFA) with the Weighted Least Squares Mean and Variance adjusted (WLSMV) estimator, which is appropriate for ordinal data. This method was selected for two reasons. First, the use of the Likert scale may be treated as interval data as long as the collected data represent composite scores from several items (Tanujaya et al., 2022). In the context of the present study, the overall PTG and PTD scales scores were treated as interval data, whereas individual item responses were treated as ordinal. Although the questionnaire used a six-point Likert scale, allowing us to initially

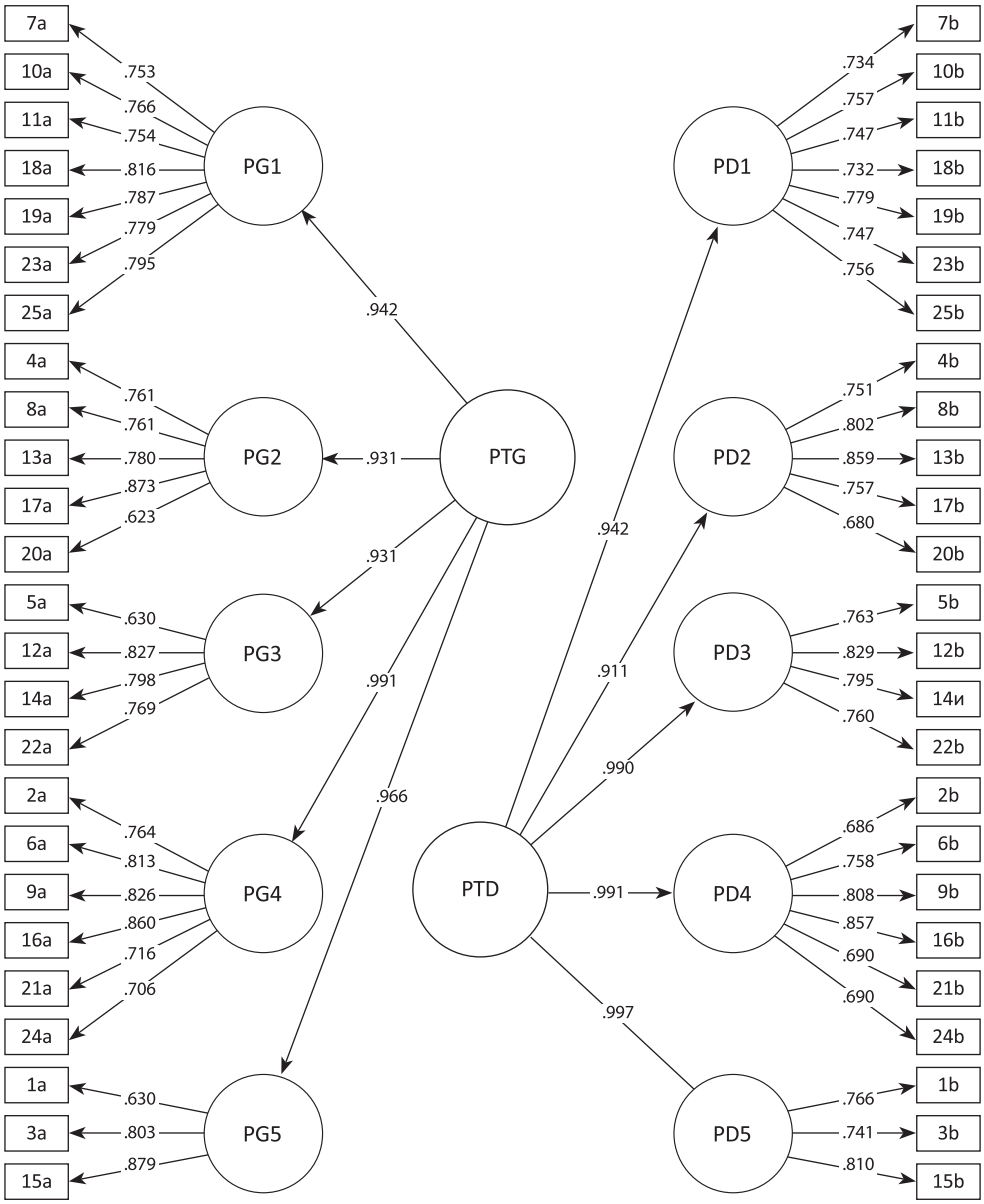

Figure 1. Final a posteriori CFA models for the PTG scale and the PTD scale.

Notes. To improve the clarity of the factor structure, the residual variances of the observed variables and first-order factors are excluded from the figure. PG 1 — Posttraumatic Growth/Relating to Others; PD 1 — Posttraumatic Depreciation/Relating to Others; PG 2 — Posttraumatic Growth/New Possibilities, PD 2 — Posttraumatic Depreciation/New Possibilities, PG 3 — Posttraumatic Growth/Personal Strength, PD 3 — Posttraumatic Depreciation/Personal Strength, PG 4 — Posttraumatic Growth/Spiritual and Existential Change, PD 4 — Posttraumatic Depreciation/Spiritual and Existential Change, PG 5 — Posttraumatic Growth/Appreciation of Life, PD 5 — Posttraumatic Depreciation/Appreciation of Life, PTG — Posttraumatic Growth, PTD — Posttraumatic Depreciation.

consider treating the data as continuous (Rhemtulla et al., 2012), the application of Mardia's test revealed that the PTG and PTD variables deviated substantially from multivariate normality (*Table 3*), including an asymmetric joint distribution and clustering of extreme values on both scales, indicating that continuous-data estimation methods were unsuitable.

Given that contemporary empirical data contribute to the conceptualization of PTG and PTD as relatively distinct phenomena (Cann et al., 2010), we examined the factor structure of the PTGDI-X separately for each corresponding dimension. The final a posteriori CFA model with standardized regression coefficients for the PTG scale and the PTD scale are presented in *Figure 1*. The obtained model fit indices for the empirical data are presented in *Table 4*.

Table 3

Results of Mardia's Test for Multivariate Normality

	Statistics	PTG	PTD
Multivariate skewness	Multivariate skewness coefficient, b1p	101.73	230.07
	Chi-square statistic, χ2	6849.84	15491.46
	P-value	<.001	<.001
Multivariate kurtosis	Observed value of the multivariate kurtosis coefficient, b2p*	879.11	1181.06
	Z-statistic	55.83	138.42
	P-value	<.001	<.001

*Note. *Expected value under normality = 675*

Table 4

Conformity Indices of Confirmatory Models

Model	CMIN	CMIN/df	NFI	TLI	CFI	RMSEA	SRMR
PTG	647.141*	2.397	.991	.994	.995	.059	.054
PTD	753.150*	2.789	.985	989	.990	.067	.072

*Notes. *p < .001; CMIN — chi-square criterion; p — significance level of chi-square; CMIN/df — chi-square divided by degrees of freedom; NFI — normed fit index, TLI — Tucker-Lewis index, CFI — comparative fit index, RMSEA — root mean square error of approximation, SRMR — standardized root mean square residual*

The regression coefficients show that the items are illustrative indicators of the questionnaire's subscales, which are also significant indicators of the total scales. The regression coefficients showed that the items loaded strongly onto their respective subscales, which were significant indicators of the overall PTG and PTD total scales.

Overall, the fit indices supported the adequacy of the constructed models for both the PTG and PTD scales. The CMIN/df ratio for both models were below 3, indicating good model fit. The incremental fit indices (NFI, TLI, and CFI) values exceeded the recommended threshold of .90, indicating satisfactory model fit. The RMSEA and SRMR values were within acceptable limits, not exceeding .08. This allows us to conclude that the factor structure of the methodology scales, identified as a result of CFA, confirms their construct validity.

PTGDI-X Scales Consistency

Given that the data deviated from a normal distribution, robust estimators of Cronbach's alpha and McDonald's omega were employed to evaluate the internal consistency of the scales (Zhang & Yuan, 2016).

Both Cronbach's alpha (PTG-total = .956; PTD-total = .942) and McDonald's omega (PTG-total = .964; PTD-total = .952) coefficients were found to be high. Likewise, the coefficients for the separate subscales indicated satisfactory internal consistency, ranging from .774 to .897, except for the Appreciation of Life subscale of the PTD, which yielded somewhat lower values, approximating .70. The obtained values generally revealed high internal consistency of the questionnaire items and demonstrated sufficient reliability of the Russian-language version of the PTGDI-X (see *Table 5*).

Table 5

The PTGDI-X Scales Internal Consistency

PTG scales	Cronbach's alpha	McDonald's omega	PTD scales	Cronbach's alpha	McDonald's omega
PTG /Relating to Others	.896	.897	PTD /Relating to Others	.825	.827
PTG /New Possibilities	.831	.838	PTD /New Possibilities	.837	.867
PTG /Personal Strength	.839	.843	PTD /Personal Strength	.825	.830
PTG /Spiritual and Existential Change	.869	.873	PTD /Spiritual and Existential Change	.802	.808
PTG /Appreciation of Life	.774	.785	PTD /Appreciation of Life	.693	.708
PTG / Total	.956	.964	PTD / Total	.942	.952

Notes. PTG - Posttraumatic Growth; PTD - Posttraumatic Depreciation.

PTGDI-X Internal Construct Validity

The results of the correlation analysis of the PTGDI-X scales are presented in the *Table 6*.

Table 6

The PTGDI-X Subscales Intercorrelations

Scales	PG 1	PD 1	PG 2	**PD 2**	**PG 3**	**PD 3**	**PG 4**	**PD 4**	**PG 5**	**PD 5**	**PGT**	**PDT**
PG 1	1											
PD 1	.26**	1										
PG 2	.74**	.35**	1									
PD 2	.23**	.70**	.16**	1								
PG 3	.74**	.25**	.77**	.12*	1							
PD 3	.22**	.77**	.22**	.77**	-	1						
PG 4	.80**	.31**	.78**	.16**	.75**	.18**	1					
PD 4	.28**	.77**	.24**	.76**	.16**	.76**	.21**	1				
PG 5	.80**	.34**	.71**	.28**	.70**	.25**	.79**	.28**	1			
PD 5	.26**	.62**	.25**	.60**	.17**	.63**	.23**	.71**	.24**	1		
PGT	.93**	.33**	.88**	.21**	.88**	.21**	.92**	.26**	.88**	.26**	1	
PDT	.31**	.89**	.30**	.90**	.20**	.88**	.27**	.90**	.34**	.76**	.32**	1

*Notes. * p<.05; ** p<.01; PG 1 - Posttraumatic Growth/Relating to Others; PD 1 - Posttraumatic Depreciation/Relating to Others; PG 2 - Posttraumatic Growth/New Possibilities, PD 2 - Posttraumatic Depreciation/New Possibilities, PG 3 - Posttraumatic Growth/Personal Strength, PD 3 - Posttraumatic Depreciation/ Personal Strength, PG 4 - Posttraumatic Growth/Spiritual and Existential Change, PD 4 - Posttraumatic Depreciation/Spiritual and Existential Change, PG 5 - Posttraumatic Growth/Appreciation of Life, PD 5 – Posttraumatic Depreciation/Appreciation of Life, PGT - Posttraumatic Growth/Total, PDT - Posttraumatic Depreciation /Total.*

PTG scales indicate multiple strong and moderate positive connections with each other. A similar structure of correlational relationships is observed across the scales of PTD dimension. Weak to moderate positive correlations were found between PTG and PTD. This indicates that the variability in one variable is accounted for only to a small extent by the variability in the other. This limits the practical significance of the association and indicates that PTG and PTD may occur together while remaining largely separate and weakly related phenomena. The structure of intercorrelations disclosed through the correlation analysis corroborates the theoretical constructs underlying the validated measure.

Nonlinear Dependence of PTG and PTD Scales

To clarify the correlation between the phenomena of the PTG and PTD scales, we used a nonlinear regression analysis. The result showed a significant inverted U-shape quadratic dependence between the total PTG scale scores and the total PTD scale scores: $F = 27.121$, $p < .001$, $\Delta R2 = .079$, $R2 = .119$ (see *Figure 2*).

The results showed that the analyzed indicators have a nonlinear dependence, i.e., an increase in the total PTG indicator has a significant moderate correlation with an increase in the total PTD indicator up to a certain level; a further increase in PTG indicators is associated with a decrease in PTD indicators.

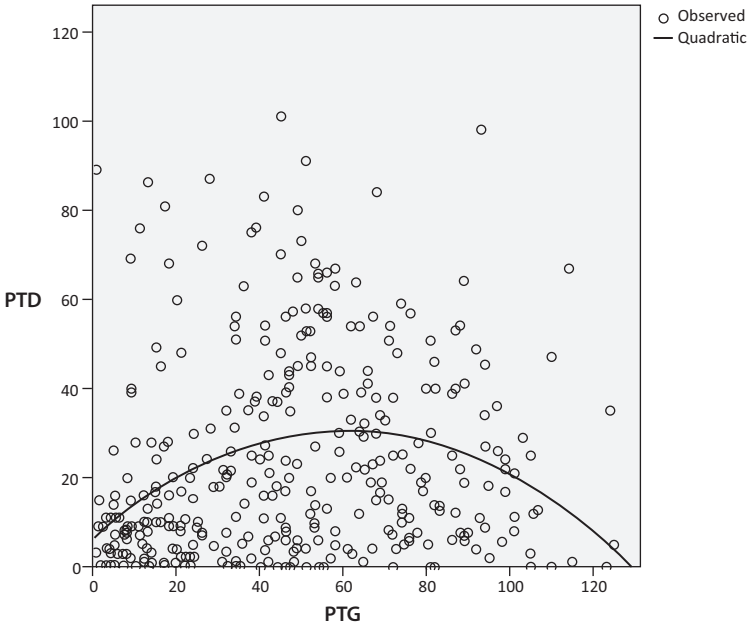

Figure 2. Dependence of the PTD level on the PTG level

Splitting the sample into two parts based on the PTG median (Md = 42) and calculating the Spearman's correlation coefficient *r* in the obtained parts showed that in the first half, the correlation is significantly moderate positive r~s~ = .467 at $p < .001$, while in the second half, the correlation is significantly weak negative r~s~ = -.147 at $p = .04$, which also indicates the presence of a nonlinear dependency between PTG and PTD indicators.

PTGDI-X External Construct Validity

Convergent and discriminant validity were evaluated by analyzing the correlations between the scales of the PTGDI-X and those of the WAS, SCL-90-R, and LREB. The findings are shown in *Table 7*.

As expected, we found significant positive correlations between the WAS subscales with the subscales of the PTG dimension, and negative correlations with the subscales of the PTD dimension. It is quite natural that beliefs in the controllability of the surrounding world, its justice, and one's own luck in life are associated with PTG in various areas. At the same time, in individuals experiencing pronounced stress and psychological traumatization, the process of depreciation is closely associated with the loss of basic trust in the world as safe, just, benevolent, and controllable.

As shown in *Table 7*, the PTGDI-X and the SCL-90-R subscales show a significant number of positive interrelations of subjectively noted symptoms of mental distress with PTG and PTD across all studied dimensions. Thus, depression and anxiety symptoms, as well as other manifestations of psychological distress, accompany both of PTG and PTD phenomena. More weak connections found between indicators of psychopathological symptomatology and PTG, compared to indicators of PTD,

Table 7

Intercorrelations Between the PTGDI-X and WAS, SCL-90-R, and LREB Variables

	PTGDI-X											
	PG 1	**PD 1**	**PG 2**	**PD 2**	**PG 3**	**PD 3**	**PG 4**	**PD 4**	**PG 5**	**PD 5**	**PGT**	**PDT**
WAS												
Benevolence of the Impersonal World	.11*	–.27**	–	–.12*	–	–.16**	–	–.14**	–	–.10*	–	–.18**
Justice	.19**	–.20**	.12*	–.18**	.19**	–.21**	.24**	–.17**	.15**	–	.20**	–.19**
Self-Worth	–	–.23**	.11*	–.29**	.16**	–.27**	–	–.29**	–	–.20**	–	–.29**
Luck	.18**	–.28**	.18**	–.35**	.26**	–.33**	.16**	–.30**	.10*	–.18**	.19**	–.32**
Self-controllability	.12*	–.19**	.22**	–.31**	.23**	–.28**	.12*	–.24**	–	–.19**	.16**	–.27**
SCL-90-R												
Somatization	.30**	.40**	.22**	.39**	.18**	.45**	.27**	.39**	.27**	.33**	.29**	.46**
Obsessive-Compulsive	.28**	.41**	.26**	.43**	.20**	.45**	.27**	.44**	.29**	.35**	.30**	.49**
Interpersonal Sensitivity	.23**	.40**	.26**	.35**	.18**	.42**	.25**	.39**	.26**	.37**	.27**	.45**
Depression	.23**	.49**	.26**	.51**	.14**	.53**	.23**	.50**	.29**	.42**	.26**	.57**
Anxiety	.24**	.41**	.25**	.40**	.14**	.45**	.25**	.43**	.27**	.35**	.26**	.48**
Hostility	.17**	.32**	.22**	.30**	.16**	.37**	.20**	.31**	.22**	.31**	.21**	.38**
Phobic Anxiety	.17**	.32**	.16**	.32**	.11*	.34**	.21**	.36**	.17**	.31**	.19**	.38**
Paranoid Ideation	.20**	.41**	.24**	.36**	.19**	.40**	.23**	.34**	.27**	.31**	.25**	.42**
Psychoticism	.25**	.45**	.25**	.42**	.17**	.47**	.27**	.42**	.27**	.36**	.27**	.50**
Global Severity Index	.30**	.48**	.29**	.48**	.20**	.52**	.30**	.49**	.33**	.42**	.32**	.56**
Symptom Distress Index	.29**	.47**	.30**	.45**	.21**	.51**	.30**	.47**	.32**	.40**	.32**	.54**
Positive Symptom Total	.18**	.39**	.17**	.43**	.11*	.44**	.16**	.42**	.20**	.38**	.19**	.49**

*Notes. * p<.05; ** p<.01; PG 1 — Posttraumatic Growth/Relating to Others; PD 1 — Posttraumatic Depreciation/Relating to Others; PG 2 — Posttraumatic Growth/New Possibilities, PD 2 — Posttraumatic Depreciation/New Possibilities, PG 3 — Posttraumatic Growth/Personal Strength, PD 3 — Posttraumatic Depreciation/Personal Strength, PG 4 — Posttraumatic Growth/Spiritual and Existential Change, PD 4 — Posttraumatic Depreciation/Spiritual and Existential Change, PG 5 — Posttraumatic Growth/Appreciation of Life, PD 5 – Posttraumatic Depreciation/Appreciation of Life, PGT — Posttraumatic Growth/Total, PDT — Posttraumatic Depreciation /Total.*

Table 8

Intercorrelations Between the PTGDI-X and WAS, SCL-90-R, and LREB Variables

	PTGDI-X											
	PG 1	**PD 1**	**PG 2**	**PD 2**	**PG 3**	**PD 3**	**PG 4**	**PD 4**	**PG 5**	**PD 5**	**PGT**	**PDT**
LREB												
Vigor	.10*	–.19**	.10*	–.21**	.20**	–.26**	.13**	–.22**	–	–.14**	.14**	–.17**
Dedication	.20**	–	.24**	–.13*	.27**	–.16**	.29**	–.11*	.21**	–	.27**	–.11*
Resource	.18**	–.13**	.20**	–.17**	.26**	–.22**	.25**	–.16**	.17**	–.10*	.24**	–.17**
Self-efficacy	.26**	–.14**	.25**	–.12*	.31**	–.18**	.25**	–.14**	.21**	–	.29**	–.14**
Exhaustion	–	.29**	–	.47**	–	.39**	–	.36**	.17**	.31**	–	.42**
Depersonalization	–	.31**	–	.38**	–	.33*	–	.33**	–	.31**	–	.39**
Destruction	.17**	.29**	.15**	.43**	–	.38**	.18**	.40**	.24**	.32**	.17**	.42**
Inefficacy	–	.27**	–.15**	.37**	–.15**	.29**	–	.30**	–	.28**	–.10*	.34**

*Notes. * p<.05; ** p<.01; PG 1 - Posttraumatic Growth/Relating to Others; PD 1 - Posttraumatic Depreciation/Relating to Others; PG 2 - Posttraumatic Growth/New Possibilities, PD 2 - Posttraumatic Depreciation/New Possibilities, PG 3 - Posttraumatic Growth/Personal Strength, PD 3 - Posttraumatic Depreciation/ Personal Strength, PG 4 - Posttraumatic Growth/Spiritual and Existential Change, PD 4 - Posttraumatic Depreciation/Spiritual and Existential Change, PG 5 - Posttraumatic Growth/Appreciation of Life, PD 5 – Posttraumatic Depreciation/Appreciation of Life, PGT - Posttraumatic Growth/Total, PDT - Posttraumatic Depreciation /Total.*

may reveal a nonlinear dependence between growth and stress symptomatology. The more pronounced strength of correlation coefficients is revealed between the overall indicator of posttraumatic depreciation (Posttraumatic Depreciation /Total, PTGDI-X), depressive symptoms, and the severity indicators of psychological distress (Depression; Global Severity Index and Symptom Distress Index, SCL-90-R).

The correlations between the scales of the LREB and PTGDI-X proved to be predictable and supported the construct validity of the questionnaire. Thus, feelings of exhaustion, cynicism, detachment, or negative attitudes toward the care recipient, as well as a sense of personal ineffectiveness in the process of caring for a seriously ill loved one, were found to be associated with the process of depreciation within ICs' value-meaning sphere. The more pronounced the devaluation, the greater the extent to which ICs are prone to destructive tension discharge, manifested in the somatization of emotional distress and addictive behaviors. The PTG scales of PTGDI-X are, as expected, positively associated with resource scales of LREB, such as Vigor, Dedication, Resources, and Self-Efficacy. Thus, the higher the level of a relative's engagement in the caregiving process, and their feelings of meaningfulness and effectiveness of their actions in overcoming the disease, the more pronounced the positive changes that occur in the relative's value system as a result of confronting a loved one's illness. It should be noted that destructive tension discharge (the Destruction scale of LREB) can accompany positive personal transformations in almost all domains of PTG.

Discussion

The aim of this study was to verify the factor structure and psychometric characteristics of the Russian-language version of the PTGDI-X. The validation procedures were carried out on a non-clinical sample of ICs caring for patients with diverse somatic, neurological, and mental disorders. The sample of ICs varied in terms of age, gender, and caregivers' family role. Most ICs were women, primarily mothers and wives, while most male caregivers were husbands. Therefore, the sample shows that those individuals are more similar to typical ICs rather than the population as a whole. As standardization of the PTGDI-X was not the primary objective of this study, greater gender and role balance in the sample wasn't our priority; this task will be addressed in our future work.

Verification of the factor structure of the analyzed version of the questionnaire, conducted through CFA, showed sufficient adequacy of the constructed models for both the PTG and PTD scales. Our findings are consistent with the results of the majority of studies (Ho et al., 2022; Ramos et al., 2016; Taku et al., 2021) and support the cross-cultural invariance of the five-factor structure of the PTGDI-X in both PTG and PTD dimensions.

It should be noted that the literature contains studies in which the five-factor model was not confirmed (Bocharov et al., 2024; Romeo et al., 2023). For example, in a study conducted on a sample of 601 Italian-speaking adults with traumatic experiences, the five-factor structure of the PTD dimension was not confirmed (Romeo et al., 2023). Moreover, deeper examinations of previous validations (*e.g.*, Baker et al., 2008; Cann et al., 2010) have consistently shown that PTD is psychometrically less stable than PTG, possibly because participants tend to respond in ways that reduce discomfort or present themselves more positively. The factor structure of PTD remains debated, with some studies supporting a unidimensional model rather than five distinct domains (Cann et al., 2010; Romeo et al., 2022). Thus, the poor fit of the PTD model in the Italian validation study is not an isolated finding, but rather a recurring challenge that warrants further investigation. The sample location and characteristics should be also considered. The Italian validation sample consisted predominantly of young, female, and highly educated participants, most of whom reported interpersonal losses (*e.g.*, the death of a relative) rather than direct life-threatening events. This sociodemographic and trauma-related profile may limit the generalizability of the findings and partly explain the weaker factor structure observed for PTD. Conversely, in our sample, the model obtained for the PTD scales is not inferior to the model obtained for PTG in terms of fit indices.

The results of the psychometric evaluation indicate a good level of reliability of the adapted tool. Cronbach's alpha and McDonald's omega coefficients were high for both the PTG-total and PTD-total scales, exceeding .90. For the separate subscales, Cronbach's alpha and McDonald's omega also demonstrated good reliability, ranging from .774 to .897, with the exception of the Appreciation of Life subscale of the PTD, where the values were slightly lower, at approximately .70. The data are consistent with the results identified during the cross-cultural adaptation of the PTGDI-X in other populations, and demonstrated the lowest consistency of statements for both the PTG and the PTD dimensions (Romeo et al., 2023).

Assessing the internal construct validity of the PTGDI-X, we assessed intercorrelations between the scale scores of PTG and PTD. These findings confirm that the phenomena of growth and depreciation can coexist in individuals following trauma (Cann et al., 2010; Pieta & Rzeszutek, 2023; Romeo et al., 2022). This, in turn, underscores the necessity of providing respondents with the opportunity to assess the degree of severity of both positive and negative consequences resulting from their exposure to psychotraumatic events.

It is important to emphasize that the nonlinear regression analysis indicates the presence of a nonlinear relationship between PTG and PTD indicators. In other words, an increase in the PTG-total scale has a significant moderate correlation with an increase in the PTD-total scale up to a certain level; a further increase in PTG indicators is associated with a decrease in PTD indicators. The data are consistent with the results obtained during the adaptation of the PTGDI-X on a Chinese sample, where a significant quadratic relationship was also found (Ho et al., 2022).

The results of the analysis of cross-convergent and discriminant validity also confirm the correspondence of the analyzed Russian-language version of the PTGDI-X to the underlying theoretical constructs. The study of correlational relationships between the PTGDI-X and WAS scales revealed that, as expected, the PTG scales show positive correlations with the scales reflecting basic trust in the world and in oneself. The results of our study are consistent with previous findings and demonstrate that belief in justice, the controllability of the surrounding world, and one's own luck in life are associated with PTG in various areas (Bayer et al., 2007; Dekel et al., 2011; Szcześniak et al., 2022). Furthermore, the present study demonstrates that the PTD scales (PTGDI-X) are negatively correlated with the scales of basic beliefs (WAS). Thus, the identified relationships indicate a close association between the processes of cognitive appraisal of the situation and the direction of personal changes in stressful or psychotraumatic circumstances (Calhoun et al., 2000). Assumptions about justice and luck, controllability of the current life situation, as well as perceived social support, act as predictors of PTG. A negative appraisal of oneself and the surrounding reality is associated with destructive consequences manifested in the form of PTD.

Our data are consistent with the results of studies reporting associations between depressive, anxiety and other psychopathological distress symptoms and the processes of PTD and PTG (Ho et al., 2022; Michélsen et al., 2017; Romeo et al., 2023). The positive link between PTG and psychopathological symptoms may reflect a facilitative role of stress symptoms in the emergence of growth, whereby initial distress triggers the cognitive processing necessary for positive transformation. Since evidence suggests that depressive and stress symptoms exhibit nonlinear relationships with PTG (Shakespeare-Finch & Lurie-Beck, 2014), the obtained data require additional clarification within the target population and may be considered a potential direction for further research.

The data obtained regarding the relationship between PTD and burnout in the process of caring for a chronically ill loved one demonstrated that under the conditions of traumatic stress, these two detrimental processes become closely entwined and mutually perpetuating, which is consistent with the results of other studies (Se-

meijn et al., 2019). PTG is positively interrelated with engagement, which suggests that the sense of meaning and the feeling of rightness regarding one's own patient care actions should be considered as a resource potential for ICs in the process of caring for a chronically ill.

Conclusion

The Russian-language version of the PTGDI-X demonstrated good psychometric characteristics. The analysis of the factor structure, conducted using CFA, indicates a sufficient correspondence of the model to the empirical data and confirms the construct validity. Evaluation of internal construct and cross validity also confirmed the theoretical constructs underlying the questionnaire.

The Russian-language version of the PTGDI-X, validated on a group of relatives providing assistance and care for seriously and/or chronically ill patients, can be used to study the phenomena of PTG and PTD that arise in a person in various stressful and psycho-traumatic situations. Employing this instrument will help consolidate information on PTG and PTD and help address current methodological and instrumental gaps in cross−cultural studies of these phenomena.

The range of the tool's application is wide, as the PTGDI-X can be used both in individual work and in clinical research. Its application will help to have a better understanding of theoretical ideas about the mechanisms of an individual's mental response to the impact of stressful or psycho-traumatic events, as well as expand clinical perspectives, allowing assessment of the individual resource potential and identification of "targets" for tailored psychotherapy.

Limitations

It should be noted that the sample is specific, which defines the essence of the stress-related burden that emerges within the context of family relationships characterized by high personal significance. This stress load is predominantly prolonged (chronic) in its course; however, it may vary significantly depending on the type of disease. The specificity of PTG/PTD may vary depending on the type and how severe and deeply felt is the experience of traumatic stress. In other words, the way traumatic stress is experienced can be influenced by the nature of the event and may differ between ICs of severely or chronically ill patients and individuals exposed to combat and natural disasters. Future research should focus on how PTG and PTD differ depending on the type of psychological stressor/traumatic impact. Future validations should include more diverse samples in term of age, gender, trauma type, and educational background to ensure the reliability of the results.

In relation to relatives of chronically ill patients, future studies will focus more on analyzing various factors, such as disease type and severity, caregivers' family role (parent, spouse, child, sibling), and their attitudes toward the patient. An additional promising area for future research is the investigation of the relationship between PTG and PTD in caregiving relatives compared with patients themselves, which may provide insight into different coping trajectories within the same stressful context.

Ethics Statement

The study was approved by the local Ethics Committee of the V.M. Bekhterev National Medical Research Center for Psychiatry and Neurology (approval No: 10/21 and No. 11/23). All participants provided written informed consent.

Author Contributions

A.Sh. and V.B. conceptualized and designed the study; T.V. analyzed the data, A.Sh. drafted the manuscript; V.B. and R.T. reviewed and edited the manuscript. All authors contributed to the final manuscript.

Conflict of Interest

The authors declare no conflict of interest.

Funding

Part of this research was supported by the Russian Science Foundation (Project No. 23-25-00090).

References

Alves, L.C.S., Monteiro, D.Q., Bento, S.R., Hayashi, V.D., de Carvalho Pelegrini, L.N., & Carvalho Vale, F.A. (2019). Burnout syndrome in informal caregivers of older adults with dementia: A systematic review. *Dementia & Neuropsychologia, 13*(4), 415–421. https://doi.org/10.1590/1980-57642018dn13-040008

Ansah, J.P., & Chiu, C.T. (2023). Projecting the chronic disease burden among the adult population in the United States using a multi-state population model. *Frontiers in Public Health, 10*, 1082183. https://doi.org/10.3389/fpubh.2022.1082183

Antonova, A.A., Yamanova, G.A., Bogovdenova, V.F., & Umarova, D.N. (2021). Osnovnye tendentsii zabolevaemosti sredi detskogo naseleniia [Main trends in morbidity among children]. *Mezhdunarodnyi Nauchno-issledovatel'skii Zhurnal* [International Research Journal], *1*(103), 6–9. https://doi.org/10.23670/IRJ.2021.103.1.054

Baker, J.M., Kelly, C., Calhoun, L.G., Cann, A., & Tedeschi, R.G. (2008). An examination of posttraumatic growth and posttraumatic depreciation: Two exploratory studies. *Journal of Loss and Trauma, 13*(5), 450–465. https://doi.org/10.1080/15325020802171367

Barello, S., Castiglioni, C., Bonanomi, A., & Graffigna, G. (2019). The Caregiving Health Engagement Scale (CHE-s): Development and initial validation of a new questionnaire for measuring family caregiver engagement in healthcare. *BMC Public Health, 19,* 1562. https://doi.org/10.1186/s12889-019-7743-8

Barrington, A., & Shakespeare-Finch, J. (2013). Posttraumatic growth and posttraumatic depreciation as predictors of psychological adjustment. *Journal of Loss and Trauma, 18*(5), 429–443. https://doi.org/10.1080/15325024.2012.714210

Bayer, S., Lev-Wiesel, R., & Amir, M. (2007). The relationship between basic assumptions, posttraumatic growth, and ambiguity tolerance in an Israeli sample of young adults: A mediation-moderation model. *Traumatology, 13*(1), 4–15. https://doi.org/10.1177/1534765607299908

Beaton, D.E., Bombardier, C., Guillemin, F., & Ferraz, M.B. (2000). Guidelines for the process of cross-cultural adaptation of self-report measures. *Spine, 25*(24), 3186–3191. https://doi.org/10.1097/00007632-200012150-00014

Bocharov, V.V., Shishkova, A.M., Tsygankova, E.S., & Chernaya, Yu.S. (2024). Faktornaia struktura Oprosnika posttravmaticheskogo rosta i posttravmaticheskogo obestsenivaniia na russkoiazychnoi vyborke rodstvennikov patsientov s tserebral'noi patologiei [Factor structure of the Posttraumatic Growth and Posttraumatic Depreciation Inventory – Expanded version on the Russian-language sample of relatives of patients with cerebral pathology]. *Pediatr* [Pediatrician], *15*(6), 93–102. https://doi.org/10.17816/PED15693-102.

Calhoun, L. G., Cann, A., Tedeschi, R. G., & McMillan, J. (2000). A correlational test of the relationship between posttraumatic growth, religion, and cognitive processing. *Journal of Traumatic Stress, 13*(3), 521–527. https://doi.org/10.1023/A:1007745627077

Cann, A., Calhoun, L.G., Tedeschi, R.G., & Solomon, D.T. (2010). Posttraumatic growth and depreciation as independent experiences and predictors of well-being. *Journal of Loss and Trauma*, *15*(3), 151–166. https://doi.org/10.1080/15325020903375826

Cejalvo, E., Marti-Vilar, M., Gisbert-Perez, J., & Badenes-Ribera, L. (2025). Stress as a risk factor for informal caregiver burden. *Healthcare, 13*(7), 731. https://doi.org/10.3390/healthcare13070731

Choi, J.Y., Lee, S.H., & Yu, S. (2024). Exploring factors influencing caregiver burden: A systematic review of family caregivers of older adults with chronic illness in local communities. *Healthcare, 12*(10), 1002. https://doi.org/10.3390/healthcare12101002

Choudry, M. M., Murray, N., Dindinger-Hill, K., Ambrose, J., Hunt, T. C., Horns, J., Martin, C., Haaland, B., Lowrance, W., Hanson, H. A., Matern, R., Cartwright, P. C., & O'Neil, B. (2024). Genitourinary cancer and family: The reverberating psychological and cardiovascular effects of a genitourinary cancer diagnosis on first-degree relatives and spouses. *Cancer, 130*(23), 4061–4070. https://doi.org/10.1002/cncr.35486

Ci Soh, X., Hartanto, A., Ling, N., Reyes, M., Sim, L., & Majeed, N.M. (2025). Prevalence of depression, anxiety, burden, burnout, and stress in informal caregivers: An umbrella review of meta-analyses. *Archives of Gerontology and Geriatrics Plus, 2*(3), 100197. https://doi.org/10.1016/j.aggp.2025.100197

Cureton, J. (2014). Trauma redefined in the DSM-5: Rationale and implications for counseling practice. *The Professional Counselor*, *4*(3), 257–271. https://doi.org/10.15241/LKJ.4.3.257

Dekel, S., Mandl, C., & Solomon, Z. (2011). Shared and unique predictors of post-traumatic growth and distress. *Journal of Clinical Psychology, 67*(3), 241–252. https://doi.org/10.1002/jclp.20747

Derogatis, L. R., Lipman, R. S., & Covi, L. (1973). SCL-90: An outpatient psychiatric rating scale-preliminary report. *Psychopharmacology Bulletin, 9*(1), 13–28.

Epstein, J., Santo, R.M., & Guillemin, F. (2015). A review of guidelines for cross-cultural adaptation of questionnaires could not bring out a consensus. *Journal of Clinical Epidemiology, 68*(4), 435–441. https://doi.org/10.1016/j.jclinepi.2014.11.021

Exenberger, S., Kumnig, M., Juen, B., Rumpold, G., & Siller, H. (2019). Dimensions of posttraumatic growth in a German-speaking sample using mixed methods. *European Journal of Psychotraumatology, 10*(1), 1701258. https://doi.org/10.1080/20008198.2019.1701258

Feriante, J., & Sharma, N. P. (2023). *Acute and Chronic Mental Health Trauma*. StatPearls Publishing.

Hacker K. (2024). The Burden of Chronic Disease. *Mayo Clinic Proceedings. Innovations, Quality & Outcomes, 8*(1), 112–119. https://doi.org/10.1016/j.mayocpiqo.2023.08.005

Ho, S. M. Y., Cheng, C. T., Shih, S. M., Taku, K., & Tedeschi, R. G. (2022). The Chinese version of Posttraumatic Growth and Depreciation Inventory-Expanded version (PTGDI-X) for cancer survivors. *Supportive Care in Cancer: Official Journal of the Multinational Association of Supportive Care in Cancer, 30*(1), 237–249. https://doi.org/10.1007/s00520-021-06223-8

Janoff-Bulman, R. (1989). Assumptive worlds and the stress of traumatic events: Applications of the schema construct. *Social Cognition*, *7*(2), 113–136. https://doi.org/10.1521/soco.1989.7.2.113

Kunz, S., Fellinghauer, C., & Peter, C. (2019). Measuring posttraumatic growth and depreciation after spinal cord injury: A Rasch analysis. *Rehabilitation Psychology, 64*(4), 407–424. https://doi.org/10.1037/rep0000288

Lim, J. M., Barlas, J., Kaur, D., & Ng, P. (2024). Unmasking the struggle: A scoping review exploring post-traumatic stress symptoms in caregivers of individuals with neurodevelopmental, psychi-

atric and neurocognitive disorders. *Trauma, Violence, & Abuse, 25*(4), 3191–3210. https://doi.org/10.1177/15248380241241018

Magomed-Eminov, M.Sh. (2008) *Fenomen ekstremal'nosti* [The phenomenon of extremity], PARF.

Manis, H. C., Allbaugh, L. J., & O'Dougherty Wright, M. (2025). Posttraumatic growth and depreciation following childhood maltreatment: Co-occurrence and cognitive differentiators. *Journal of Aggression, Maltreatment & Trauma, 34*(3), 367–386. https://doi.org/10.1080/10926771.2025.2473460

Michelsen, H, Therup-Svedenlof, C, Backheden, M, & Schulman, A. (2017). Posttraumatic growth and depreciation six years after the 2004 tsunami. *European Journal of Psychotraumatology, 8*(1), 1302691. https://doi.org/10.1080/20008198.2017.1302691

Moreta-Herrera, R., Regatto-Bonifaz, J., Viteri-Miranda, V., López-Barrionuevo, C., Alejandro, R. J., & Moreta, M. E. (2025). Impact of post-traumatic stress associated with COVID-19 after two years of the pandemic: Differences in latent means in Ecuadorian adults. *Hellenic Journal of Psychology, 22*(2), 208–223. https://doi.org/10.26262/hjp.v22i2.10390

Oshiro, R., Kopitz, J., Soejima, T., Kibi, S., Kamibeppu, K., Sakamoto, S., & Taku, K. (2019). Perceptions of positive and negative changes for posttraumatic growth and depreciation: Judgments from Japanese undergraduates. *Personality and Individual Differences, 137*, 17–21. https://doi.org/10.1016/j.paid.2018.07.040

Padun, M.A., & Kotel'nikova, A.V. (2008). Modifikatsiia metodiki issledovaniia bazisnykh ubezhdenii lichnosti R. Ianoff-Bul'man [Modification of R. Janoff-Bulman's methodology for the study of basic personality beliefs]. *Psikhologicheskii Zhurnal* [Psychological Journal], *29*(4), 98–106.

Palacio, C., & Limonero, J.T. (2020). The relationship between the positive aspects of caring and the personal growth of caregivers of patients with advanced oncological illness: Posttraumatic growth and caregiver. *Support Care Cancer, 28*(7), 3007–3013. https://doi.org/10.1007/s00520-019-05139-8

Pieta, M., & Rzeszutek, M. (2023). Posttraumatic growth and posttraumatic depreciation among people living with HIV: The role of resilience and HIV/AIDS stigma in the person-centered approach. *AIDS Care, 35*(2), 230–237. https://doi.org/10.1080/09540121.2022.2141184

Platte, S., Wiesmann, U., Tedeschi, R. G., Taku, K., & Kehl, D. (2023). A short form of the posttraumatic growth and posttraumatic depreciation inventory - expanded (PTGDI-X-SF) among German adults. *Psychological Trauma: Theory, Research, Practice, and Policy, 15*(5), 838–845. https://doi.org/10.1037/tra0001378

R Core Team. (2026). *R: A language and environment for statistical computing* (Version 4.6.0) [Computer software]. R Foundation for Statistical Computing. https://doi.org/10.32614/R.manuals

Ramos, C., Leal, I., Maroco, A.L., & Tedeschi, R.G. (2016). The posttraumatic growth inventory: factor structure and invariance in a sample of breast cancer patients and in a nonclinical sample. *The Spanish Journal of Psychology,19*, E64. https://doi.org/10.1017/sjp.2016.65.

Rhemtulla, M., Brosseau-Liard, P. É., & Savalei, V. (2012). When can categorical variables be treated as continuous? A comparison of robust continuous and categorical SEM estimation methods under suboptimal conditions. *Psychological Methods, 17*(3), 354–373. https://doi.org/10.1037/a0029315

Romeo, A., Di Tella, M., Rutto, F, Castelli, L.1, Taku K., & Loera, B.L. (2023). Expanded version of the Posttraumatic Growth and Depreciation Inventory: Scale validation and refinement among Italian adults. *Psychological Trauma, 15*(1), 1–9. https://doi.org/10.1037/tra0001203

Romeo, A., Castelli, L., Zara, G., & Di Tella, M. (2022). Posttraumatic Growth and Posttraumatic Depreciation: Associations with Core Beliefs and Rumination. *International Journal of Environmental Research and Public Health, 19*(23), 15938. https://doi.org/10.3390/ ijerph192315938

Rosseel, Y. (2012). lavaan: An R Package for Structural Equation Modeling. *Journal of Statistical Software, 48*(2), 1–36. https://doi.org/10.18637/jss.v048.i02

Semeijn, J., van Ruysseveldt, J., Vonk, G., & van Vuuren, T. (2019). In flight again with wings that were once broken; Effects of post-traumatic growth and personal resources on burnout recovery. *International Journal of Workplace Health Management, 12*(5), 387–403. https://doi.org/10.1108/IJWHM-01-2019-0006

Sergienko, A.I. (2019). Emotsional'naia dezadaptatsiia i posttravmaticheskii rost roditelia rebenka s OVZ: sviaz' s sub'ektivnym vospriiatiem tiazhesti simptomatiki i s vozrastom rebenka [Emotion-

al disadaptation and post-traumatic growth in parents of children with disabilities: A link with the subjective perception of the severity of symptoms and the age of the child]. *Klinicheskaia i Spetsial'naia Psikhologiia* [Clinical and Special Psychology], *8*(4), 93–106. https://doi.org/10.17759/cpse.2019080406

Shakespeare-Finch, J & Lurie-Beck J. (2014). A meta-analytic clarification of the relationship between posttraumatic growth and symptoms of posttraumatic stress disorder. *Journal of Anxiety Disorder, 28*(2), 223–229. https://doi.org/10.1016/j.janxdis.2013.10.005

Slater, P., & Hasson, F. (2025). Quantitative research designs, hierarchy of evidence and validity. *Journal of Psychiatric and Mental Health Nursing, 32*(3), 656–660. https://doi.org/10.1111/jpm.13135

Shishkova, A., & Bocharov, V. (2021). *Emotsional'noe vygoranie rodstvennikov bol'nykh s khimicheskoi zavisimost'iu: metodologiia i instrumentarii otsenki* [Emotional burnout of relatives of patients with chemical addiction: Methodology and assessment]. Nestor- History.

Shishkova, A., Bocharov, V., Strizhitskaya, O., & Vuks, A. (2021). Psikhometricheskie kharakteristiki metodiki "Uroven' emotsional'nogo vygoraniia rodstvennikov" [Psychometric properties of the "Level of Relatives' Emotional Burnout" Questionnaire]. *Konsul'tativnaia Psikhologiia i Psikhoterapiia* [Counseling Psychology and Psychotherapy], *29*(2), 100–118. https://doi.org/10.17759/cpp.2021290205

Shishkova, A.M., Bocharov, V.V., Tsygankova, E.S., & Chernaya, Yu.S. (2024). Kross-kul'turnaia adaptatsiia i pervichnaia aprobatsiia russkoiazychnoi versii metodiki Oprosnik Posttravmaticheskogo Rosta i Posttravmaticheskogo Obestsenivaniia—OPRPO [Cross-cultural adaptation and primary approbation of the Russian version of Posttraumatic Growth and Depreciation Inventory – Expanded version]. *Obozrenie Psikhiatrii i Meditsinskoi Psikhologii Imeni V.M. Bekhtereva* [V.M. Bekhterev Review of Psychiatry and Medical Psychology], *58*(3), 74–85. http//doi.org/10.31363/2313-7053-2024-942

Steidl, A., Kamara, A., Aziz, A., & Exenberger, S. (2024). Posttraumatic growth, posttraumatic depreciation and beyond: Exploring the applicability of quantitative measures in Sierra Leone. *European Journal of Trauma & Dissociation, 8*(2), 100401. https://doi.org/10.1016/j.ejtd.2024.100401

Szcześniak, M., Madej, D., & Bielecka, G. (2022). Relationship between world assumptions and posttraumatic growth among polish cancer patients: Moderating effect of rumination. *International Journal of Environmental Research and Public Health, 19*(19), 12444. https://doi.org/10.3390/ijerph191912444

Taku, K., Tedeschi, R. G., Shakespeare-Finch, J., Krosch, D., David, G., Kehl, D., Grunwald, S., Romeo, A., Di Tella, M., Kamibeppu, K., Soejima, T., Hiraki, K., Volgin, R., Dhakal, S., Zięba, M., Ramos, C., Nunes, R., Leal, I., Gouveia, P., ... Calhoun, L. G. (2021). Posttraumatic growth (PTG) and posttraumatic depreciation (PTD) across ten countries: Global validation of the PTG-PTD theoretical model. *Personality and Individual Differences, 169*, 110222. https://doi.org/10.1016/j.paid.2020.110222

Tamarana, R., Jain, E., Sinha, S., Manjunatha, S., Amissah, A.A., & Balaji, B. (2023). Development and initial evaluation of Informal Caregiver Burnout Scale. *Indian Journal of Continuing Nursing Education*, *24*(1), 87–91. https://doi.org/10.4103/ijcn.ijcn_83_22

Tanujaya, B., Prahmana, R. C. I., & Mumu, J. (2022). Likert scale in social sciences research: Problems and difficulties. *FWU Journal of Social Sciences, 16*(4), 89–101. https://doi.org/10.51709/19951272/Winter2022/7

Tarabrina, N.V. (2001). *Praktikum po psikhologii posttravmaticheskogo stressa* [Workshop on the psychology of post-traumatic stress]. Piter.

Tedeschi, R.G., & Calhoun, L.G. (1996). The Posttraumatic Growth Inventory: Measuring the positive legacy of trauma. *Journal of Traumatic Stress, 9*(3), 455–471. https://doi.org/ 10.1007/BF02103658

Tedeschi, R. G., & Calhoun, L. G. (2004). Target article: "Posttraumatic growth: Conceptual foundations and empirical evidence". *Psychological Inquiry, 15*(1), 1–18. https://doi.org/10.1207/s15327965pli1501_01

Tedeschi, R. G., Cann, A., Taku, K., Senol-Durak, E., & Calhoun, L. G. (2017). The Posttraumatic Growth Inventory: A revision integrating existential and spiritual change. *Journal of Traumatic Stress, 30*(1), 11-18. https://doi.org/10.1002/jts.22155.

Weiss, T., & Berger, R. (2010). Posttraumatic growth around the globe: Research findings and practice implications. In T. Weiss & R. Berger (Eds.), *Posttraumatic growth and culturally competent practice: Lessons learned from around the globe* (pp. 189–195). John Wiley & Sons, Inc.

yWorks GmbH. (2023). yEd Graph Editor (Version 3.23.2) [Computer software]. https://www.yworks.com/products/yed

Zhang, Z., & Yuan, K. H. (2016). Robust coefficients Alpha and Omega and Confidence Intervals with outlying observations and missing data: Methods and software. *Educational and Psychological Measurement*, *76*(3), 387–411. https://doi.org/10.1177/0013164415594658

Zhong, M., Sun, S., Long, J., Yuan, M., Wang, M., & Zhang, Z. (2024). Tools to measure the burden on informal caregivers of cancer patients: A literature review. *Journal of Clinical Nursing, 33*(8), 2949–2970. https://doi.org/10.1111/jocn.17112

Original manuscript received April 23, 2026
Revised manuscript accepted August 27, 2026
First published online September 1, 2026

To cite this article: Shishkova, A.M., Bocharov, V.V., Valieva, T.V., Tedeschi, R.G. (2026). Psychometric Characteristics of the Russian-Language Version of the Posttraumatic Growth and Depreciation Inventory — Expanded Version (PTGDI-X). *Psychology in Russia: State of the Art*, *19*(3), 106–130. DOI: 10.11621/pir.2026.0306

Psychology in Russia: State of the Art
Volume 19, Issue 3, 2026

Russian Psychological Society
Lomonosov Moscow State University

PERSONALITY PSYCHOLOGY

The Values System of Russian University Students Compared with Their Level of Aggression

Ivan A. Konovalov[a*], Anton L. Linkov[b], Artur A. Rean[a],

[a] *Federal Scientific Center for Psychological and Interdisciplinary Research, Moscow, Russian Federation*
[b] *Moscow Aviation Institute (National Research University), Russian Federation*

*Corresponding author. E-mail: iv.konovalov@yandex.ru

Background. The study of the connections between values and the aggression of student youth is an independent area of scientific interest in social psychology and has significant implications for youth policy.

Objective. To study the values of Russian university students based on Schwartz's values theory, and to compare those values with their level of aggression.

Design. The sample consisted of 2,328 university students from nine cities in the Russian Federation. The data was collected online using the Yandex Forms platform. Multi-site convenience sampling with gatekeeper-mediated recruitment was employed. The PVQ-RR instrument was used to evaluate the significance of different values in accordance with Schwartz's framework. Aggression indicators were assessed using the Buss-Perry questionnaire. Data analysis was performed in RStudio. Student's t-test and Mann-Whitney U test were used in repeated trials.

Results. Respondents with high physical aggression have significantly higher levels of Stimulation and Power. Respondents with low physical aggression have significantly greater values of Security, Conformity, Humility, and Universalism. High anger relates to values of Self-Direction — Action, Stimulation, Hedonism, Achievement, Power, and Face. A low level of anger is associated with the values Tradition, Conformity, Humility, Universalism, and Benevolence — Dependability. Low hostility corresponds with the expression of Universalism — Tolerance.

Conclusion. Within a sample of Russian students a preference for personal-focused values is associated with higher levels of aggression, while a preference for social-focused values is linked to lower levels.

Keywords: aggression, anger, hostility, values, Russia student youth

© Konovalov, I.A., Linkov, A.L., Rean, A.A., 2026

ISSN 2074-6857 (Print) / ISSN 2307-2202 (Online)
http://psychologyinrussia.com

Introduction

The relationship between values and aggressive behavior in certain social groups is a fundamental socio-psychological issue and an urgent applied task related to assessing risk of deviance and social instability, especially when the social group in question is students.

Schwartz's Values Theory defines values as "desirable, trans-situational goals, varying in importance, that serve as guiding principles in people's lives" (Schwartz, 2006, p. 1). In the theory of value orientations, Schwartz identifies 10 values: Achievement, Power, Stimulation, Hedonism, Universalism, Self-Direction, Security, Tradition, Benevolence, and Conformity. Each value is located on a continuum, forming an interconnected system consisting of interdependent elements. That is, they form a dynamic balance: the strengthening of one can occur due to the weakening of the opposite.

These values, in turn, are grouped into meta-values: Conservation (which includes the values of Tradition, Conformity, and Security), Openness to change (Self-Direction and Stimulation), Self-Enhancement (Power and Achievement), and Self-Transcendence (Universalism and Benevolence). The value of Hedonism is included in both the Openness to change and Self-Enhancement groups. Meta-values consist of two main categories ("focuses"): personal values (such as Openness to Change and Self-Enhancement) and social values (Self-Transcendence and Conservation). Later, Schwartz (2022) proposed a revised and updated concept that includes 19 values. This new concept divides the value of Universalism into three subcategories: Universalism — Concern, Universalism — Tolerance, and Universalism — Nature. Benevolence has been divided into Benevolence — Caring and Benevolence — Dependability, while Self-Direction has two dimensions: Self-Direction — Thought and Self-Direction — Action. Conformity, on the other hand, has been split into Conformity — Rules and Conformity — Interpersonal. Finally, Power has been separated into Power — Dominance and Power — Resources. Other values have remained unchanged.

Turning to the second variable of the present study, aggression, a classical definition was proposed by Buss (1961, p. 1) as "a response that delivers noxious stimuli to another organism". However, from a critical perspective, Krahe (2013, p. 19) noted that "this purely behaviourist definition is too broad in some ways and too narrow in others". A more conventional definition was offered by Baron & Richardson (1994, p. 7): "any form of behavior directed toward the goal of harming or injuring another living being who is motivated to avoid such treatment". The current research adopts the multidimensional construct of aggression developed by Buss & Perry (1992), which encompasses physical aggression, verbal aggression, anger, and hostility. It is important to note that in the Russian adaptation of this construct by Enikolopov & Tsibulsky (2007), three core dimensions were identified—physical aggression, anger, and hostility—along with four scales, including an overall aggression score. The present study focuses on these three fundamental dimensions of aggression.

This study examines the relationship between personal values and aggression. The theoretical review considers various indicators of aggression and deviant behavior (without conceptual specification) that have been studied in relation to Schwartz's

value framework. Notably, prior research has revealed significant associations between aggression and both the Self-Transcendence and Self-Enhancement domains within Schwartz's theory; these higher-order value domains are examined in both the theoretical review and the empirical analysis of the present study.

An analysis of the literature reveals significant connections between distinct value orientations and aggression indicators. In particular, recent studies have demonstrated links between aggression and two higher-order value domains: Openness to change, which comprises Self-Direction and Stimulation, and Self-Enhancement, which includes Power and Achievement.

For example, in the study by Obukhovskaya and Konopleva (2023), Self-Direction and Security were found to be the most significant values among adolescents enrolled in schools and in the KDN (Commission on Juvenile Affairs), while Tradition was the least important value in both groups. The group of teenagers registered in the KDN had a higher level of overall aggression. Differences in the hierarchical order of values based on priority were observed only for Universalism in the control group and Benevolence among those registered in the KDN, with both values ranking as the third most significant in the hierarchical structures of both groups. Considering values according to the methodology used by Schwartz, the adaptation of Karandashev's study found that, at the level of individual priorities, the values of Self-Direction, Hedonism, and Universalism were most significant in the control group, while Tradition, Conformity, and Power were least important. In the group registered in the KDN, Self-Direction, Benevolence, and Stimulation were the most important values, while Universalism, Power, and Conformity were the least significant. Thus, there is a certain degree of similarity in the hierarchical structure of values between the two groups. The results of this study suggest that hierarchical value structure is not a clear predictor of aggressive behavior, as many studies have found a positive correlation between values such as Power and level of aggression (McDonald et al. 2015; Rubel-Lifschitz et al., 2021). However, this correlation needs to be further investigated to fully understand the relationship between these variables .

It is important to note that another study (Baeva et al., 2022) showed that adolescents with high and low levels of aggression have some differences in their priority values. Specifically, pronounced differences can be seen for the values Hedonism and Security, as well as similar values such as Conformity and Stimulation. Values such as Stimulation and Power are more significant in the risk group, whereas Security is more significant in control groups. The sample size in the first study was 40, whereas the sample in the second study was 2,647, which increases the reliability of results. However, both studies agreed that Security remains a priority for control groups, while Stimulation remains a priority in risk groups. This aspect warrants further analysis, particularly regarding the Schwartzian value of security. As a core component of the higher-order Conservation dimension, this construct implies a desire for stability, order, and personal as well as collective safety (Schwartz, 1992, as cited in Karandshev, 2004). Note that in the PVQ questionnaire version, this is a single value, whereas in the PVQ-RR questionnaire, it is divided into two values: Security — Personal and Security — Societal (Schwartz et al., 2012). The value of Stimulation in Schwartz's theory belongs to the Openness to Change block and reflects a desire for

novelty, risk, searching for strong sensations, and exciting experiences. Individuals with a high level of Stimulation tend to be adventurous, experimental, and avoid routine.

It is also worth taking a closer look at the relationship between the values of Stimulation and Power with aggressive behavior. For example, a German study conducted by Seddig and Davidov (2018) examined the relationship between basic human values, attitudes towards violence, and violent behavior among young people in Germany, with a final sample size of 1,810 participants. This study showed that the meta-values of Self-Transcendence (Universalism and Benevolence), Conservation (Conformity and Tradition), and Security are negatively associated with violent behavior, while Power and Stimulation are positively associated. However, it is worth noting that although individual values may not always correlate with aggression, when combined into meta-value categories, they can reveal new patterns (Schwartz, et al. 2012). Also, in the study by Salakhova et al. (2016), in a sample of 98 people, it was found that the value of Stimulation was significantly higher among young people who were prone to auto-aggressive behavior than in those who did not demonstrate such tendencies. Also, in the study by Oshchepkov, Piskunova, and Salakhova (2023), students with greater propensity for aggressive behavior had higher levels of Achievement and Power.

Empirical evidence regarding this relationship remains inconsistent. For example, Khusainova and Khakimzyanova (2019) observed that among 100 adolescents with deviant behavior, the prioritization of Stimulation was lower and Security was higher than in their non-deviant peers, despite the former group demonstrating higher physical aggression scores. In a parallel investigation, Usacheva and Parshutin (2023) found that a cohort of cadets with low aggression profiles was predominantly guided by the values of Power, Benevolence, and Achievement. These empirical contradictions can likely be ascribed to specific sample characteristics (namely, sample size, socio-demographic composition, and assessment methodologies) or to macro-level social context variables, including the degree of social tension, dominant societal norms, and the widespread adoption of aggressive interaction strategies.

Examining the intersection of social values and aggression reveals the distinct roles of Universalism and Benevolence. Within Schwartz's theory, Universalism belongs to the Self-Transcendence dimension, characterized by the understanding, acceptance, and protection of the welfare of all people and nature. While conceptually adjacent to Benevolence, the latter is confined to close interpersonal relationships, whereas Universalism encompasses the broader society and the natural environment. Consequently, these values are theoretically opposed to those emphasizing personal ambition and dominance — namely Power and Achievement, which focus on social superiority and the pursuit of self-interest.

A study by Benish-Weisman (2015) found a negative association between Universalism and aggression, suggesting that individuals who value Universalism may be less likely to engage in aggressive behavior. This means that teenagers with strong universalist values are less likely to engage in aggressive behavior. A study by Knafo et al. (2008) found a positive association between the value of Power and aggression, while Universalism and Conformity showed a negative correlation. Similarly, a study

by Seddig and Davidov (2018) found that the meta-values of Self-Transcendence (Universalism and Benevolence) have a negative association with violent behavior. However, in a study by Chudova (2014), people who preferred the meta-value of Self-Transcendence were found to have increased levels of hostility, anger, and physical aggression.

Shifting focus to the Conservation domain, the relationship between Conformity and aggression also presents a significant area of inquiry. For example, in a study by Meshkova (2018), it was shown that in situations of anonymous revenge, deviant behaviors can be mediated by the reduced significance of Conformity. Similar findings were also reported in a study by Knafo (2008), which found a correlation between Conformity and violent behavior. In other words, individuals who value Conformity are less likely to display aggressive behavior, as confirmed by several studies (Myburgh et al., 2020; Velensia et al., 2021). It should be noted that this relationship can also be observed in the context of personal dispositions related to aggression. For example, in a study by Youn & Lee (2016), on a sample of 633 participants, it was found that the level of Conformity varies among groups with different levels of psychopathic traits. The control group demonstrated a higher level of Conformity compared to groups with primary and secondary psychopathy.

Based on the studies reviewed, it seems that the values of Power and Stimulation, which are included in the personal focus, are associated with a higher level of aggression. On the other hand, the values of Universalism, Conformity, and Benevolence, which are part of the social focus, are associated with lower levels of aggression.

Beyond the aforementioned factors, gender constitutes a critical variable in understanding the link between values and aggression. Several studies have shown (Bardi & Schwartz, 2003; Benish-Weismann, Daniel, & Knafo-Noam, 2017; Sagiv & Schwartz, 2022) that the values of boys and girls can have a different impact on aggression. For instance, a study by Benish-Weissmann and McDonald (2015) examined the link between values, gender, and aggression in a sample of 642 adolescents. The study found that among girls, there was no positive association between the meta-value Self-Enhancement (which includes values like Achievement and Power) and aggression. However, this was not the case for boys. According to the study, there was no significant association between the Openness to Change meta-value (Self-Direction, Stimulation, and Hedonism) and aggression among girls. However, among boys, Openness to Change was positively associated with aggression.

For boys, there was no association between values from the Self-Transcendence meta-category (Universalism and Benevolence) and aggression, when self-awareness was low. Boys with high self-awareness showed a negative relationship between Self-Transcendence values and aggression. Girls with low self-awareness also showed a negative association, but girls with high self-awareness did not show a significant association. Considering the studies, we've cited, it's important to emphasize the significance of gender when studying the interaction between values and aggression.

Summarizing the above, we can draw several conclusions:

1) The relationship between values and aggression is not linear. Some values, such as Power and Stimulation, are often associated with a higher level of

aggression, while others, like Universalism and Benevolence, are associated with a lower level (*e.g.*, Benish-Weismann, Daniel, & Knafo-Noam, 2017), various studies show conflicting results (*e.g.*, Chudova, 2014). This could be due to differences in samples, research methods, or the social context.

2) There is a difference in focus. Values related to personal goals, such as Power and Stimulation, are positively correlated with aggression. Values related to social goals, like Universalism, Conformity, and Benevolence, are negatively correlated. This pattern is supported by both Russian and international studies (*e.g.*, Meshkova, 2018; Seddig & Davidov, 2018).
3) Gender plays a role. Some studies have found that women tend to value social harmony and cooperation more than men, which may explain why they are less likely to engage in aggressive behavior. However, this is not a universal finding, and more research is needed to fully understand the complex relationship between gender and aggression. Sagiv & Schwartz (2022) show that in boys and girls, values can be associated with aggression in different ways. In boys, for example, Self-Enhancement values, such as Achievement and Power, are significantly linked to aggression, whereas in girls, this relationship does not exist. This underscores the importance of considering gender differences when analyzing the value predictors of aggressive behavior.

Recent research has highlighted the critical role of values in aggression. It is well established that an individual's value hierarchy significantly influences aggressive tendencies. A growing body of literature has demonstrated that self-enhancement values (*e.g.*, Power, Stimulation) are positively associated with aggression, while socially oriented values (*e.g.*, Universalism, Benevolence) are negatively associated, although overall findings remain complex and sometimes conflicting. Considerable attention has been devoted to understanding these non-linear dynamics and gender-specific pathways, such as the male-specific link between self-enhancement and aggression. However, little is known about whether these effects hold across different populations, such as Russian university students. This is particularly relevant given that the present study addresses the transition from the value-normative uncertainty that characterized Russian youth in the 1990s and early 2000s (Sobkin & Kalashnikova, 2021) to the contemporary context, where the deliberate formation of youth values has become a central objective of national youth policy. Addressing this gap is crucial for advancing theoretical models of value-behavior congruence across diverse cultural contexts. A key distinction of the present study lies in its methodological approach to assessing differences in aggression indicators across Schwartz's value dimensions. Specifically, we examine distinct facets of aggression — namely physical aggression, anger, and hostility — within this specific cultural context. Methodologically, this research employs an extreme groups approach in conjunction with bootstrapping procedures. The aim is to evaluate the stability of differences between groups exhibiting pronounced versus attenuated expression of specific values, relative to the predominant value pattern within the sample.

The current study was guided by two primary objectives:

1. To study the values of Russian students based on Schwartz's values theory.
2. To compare the values of Russian students with their Aggression levels.

Methods

Participants

The sample consisted of 2,328 university students from nine cities in the Russian Federation. The average age of respondents was 20.9 years (*standard deviation* = 2.94), and the median age was 20 years. Most respondents (83%) were female. Approximately 52% of the students had a pedagogical or psychological field of study, 23% had a social and humanities field, 9% had a natural science field, and 7% had a technical field. Therefore, the sample can be considered representative of students in psychological, pedagogical, and socio-humanities fields of study in Russia. The data was collected online using the Yandex Forms platform.

Data were collected using a non-probability, multi-site convenience sampling design with gatekeeper-mediated recruitment. The research team contacted coordinators at nine partner universities with a request to recruit up to 500 respondents per institution. Coordinators used their own communication channels (email lists, messaging apps, social media) to distribute the survey link.

The final sample comprised 2,338 completed responses. Due to the recruitment design (gatekeeper-mediated dissemination) and platform limitations (only fully completed surveys were recorded; partial responses and click-through data were not tracked), calculation of a standard response rate following AAPOR guidelines was not feasible, as the denominator (number of individuals actually invited) remains unknown.

Under a conservative assumption that each coordinator invited the target number of 500 individuals, the estimated response rate is approximately 52.0% (2,338 ÷ 4,500).

Questionnaires

The PVQ-RR (Schwartz et al., 2012) instrument was employed to evaluate the significance of different values in accordance with Schwartz's framework. Aggression indicators were assessed using the Buss-Perry questionnaire (Buss & Perry, 1992; Enikolopov & Tsibulsky, 2007).

The Schwartz values questionnaire (PVQ-RR) comprised 57 items across 19 dimensions, while the Buss-Perry Aggression Questionnaire included 29 items divided into four subscales. Both instruments utilized Likert-type response formats (6-point and 5-point, respectively), with dimension scores calculated as the mean of the respective items.

Table 1 contains the main descriptive statistics and reliability indicators (Cronbach's alpha and McDonald's omega) for scales used in the research.

According to *Table 1*, the reliability indicators for the most scales are within an acceptable range. Although the internal consistency coefficients for five subscales ranged from .60 to .70 (and .50 to .60 for two scales), these estimates are lower than .70, but could be considered acceptable for exploratory research purposes, provided that interpretations are made at the group level.

Table 1

Descriptive Statistics of Variables and Reliability Indicators

Variables	n	M	SD	Mdn	Sk	Ku	SE	α	ω
Self-Direction — Thought	2,335	11.57	1.99	12	–.81	2.14	.04	.55	.59
Self-Direction — Action	2,336	11.47	2.16	12	–.63	1.39	.04	.65	.66
Stimulation	2,336	9.59	2.34	10	–.12	.34	.05	.51	.55
Hedonism	2,336	11.16	2.27	11	–.63	1.17	.05	.63	.65
Achievement	2,336	10.88	2.40	11	–.48	.59	.05	.63	.67
Power — Dominance	2,336	6.92	3.13	7	.11	–.21	.06	.77	.78
Power — Resources	2,336	9.37	2.73	9	–.36	.33	.06	.66	.71
Face	2,336	10.94	2.67	11	–.67	.78	.06	.79	.79
Security — Personal	2,336	11.62	2.29	12	–.92	1.79	.05	.65	.66
Security — Societal	2,336	11.14	2.61	11	–.92	1.43	.05	.73	.76
Tradition	2,336	8.78	3.29	9	–.39	–.08	.07	.82	.83
Conformity — Rules	2,336	10.01	2.72	10	–.71	1.15	.06	.79	.79
Conformity — Interpersonal	2,336	9.41	2.79	10	–.44	.43	.06	.73	.75
Humility	2,335	9.25	2.60	9	–.55	.77	.05	.61	.63
Universalism — Concern	2,336	11.24	2.37	11	–.83	1.58	.05	.70	.72
Universalism — Nature	2,336	9.74	2.66	10	–.40	.61	.06	.77	.79
Universalism — Tolerance	2,330	10.71	2.49	11	–.64	1.00	.05	.68	.69
Benevolence — Caring	2,336	12.39	2.27	13	–1.20	2.61	.05	.79	.80
Benevolence — Dependability	2,336	12.14	2.22	12	–1.11	2.77	.05	.75	.76
Physical aggression	2,334	17.72	6.38	16	.97	.75	.13	.81	.85
Anger	2,336	18.86	6.37	19	.29	–.47	.13	.83	.86
Hostility	2,336	19.51	6.42	19	.37	–.16	.13	.78	.84

Note. n = Number of respondents. M = Mean. SD = Standard deviation. Mdn = Median. Sk = Skew. Ku = Kurtosis. SE = Standard Error. α = Cronbach's α. ω = McDonald's ω.

Procedure

In the first stage, the comparability of the groups was assessed. The groups with low levels of aggression were formed using the formula: *M-SD*. The groups with high levels of aggression were formed using the formula: *M+SD*. This technique was consistent with the scale norms described in (Enikolopov & Tsibulsky, 2007). There were 161 respondents (64.7% — female) in the high group for physical aggression and 1,014 (92.7% — female) in the low group. There were 185 respondents (87.5% — female) in the high group for anger and 884 (79.9 % — female) in the low group. There were 202 respondents (87.1 % — female) in the high group for hostility and 810 (81.5% — female) in the low group.

Among respondents classified as High on Physical Aggression, only 31.4% ($n = 50$) were also High on Anger, while the majority (64.2%, $n = 102$) fell into the Mid range on Anger, and only 4.4% ($n = 7$) were Low. Similarly, only 32.7% of High Physical respondents were also High on Hostility, and 35.9% of High Anger respondents were also High on Hostility. Pearson correlations between continuous subscale scores ranged from $r = .41$ to $r = .58$ (all $p < .001$), indicating moderate associations. The results of overlap analysis are presented in Appendix A.

In the second stage, the goal was to assess the differences in Schwartz's values among respondents with low and high levels of aggression. The extreme groups method was used. Since the differences in group sizes were significant, it was decided to implement a multiple group equalization procedure followed by an assessment of the differences. Bootstrapping (a Monte Carlo simulation with resampling) (Efron & Tibshirani, 1994) was used to solve this task.

The procedure was implemented as follows: a random subsample with a comparable number of respondents was repeatedly extracted from the group with a larger number of respondents, and the differences were evaluated using the Student's *t*-test and the Mann-Whitney *U*-test, and their consistency was assessed according to the recommendations (Korneev & Krichevets, 2011). The effect size was calculated according to the recommendations (Cohen, 2013). At each repetition, a subsample of 200 respondents was extracted from the larger sample (the group with low levels of physical aggression, anger, and hostility). The procedure was repeated 100 times for each comparison. Reproducing significant differences in 95% of cases was considered a stable result. A correction for multiple comparisons using the Benjamini-Hochberg method was used.

Given that the objective was to evaluate differences using the extreme groups approach, we employed standard statistical tests for comparing means, namely the independent-samples *t*-test and the Mann-Whitney *U* test. Taking into account the methodological limitations of each test, as detailed by Korneev and Krichevets (2011), we applied both indices and assessed their consistency, in accordance with the recommendations outlined in their study. Cohen's *d* was used to estimate the effect size, calculated based on the *t*-test results. Because a resampling procedure was conducted, the effect size estimates in *Tables 2–4* are presented as means with 95% CI.

Data analysis was performed in RStudio (R version 4.3.1; R Core Team, 2023; Posit Team, 2024). Data manipulation and processing were conducted using readxl (Wickham & Bryan, 2023), dplyr (Wickham et al., 2023), purrr (Wickham & Henry, 2023), and sjmisc (Ludecke, 2023). Statistical calculations and descriptive analyses were performed using psych (Revelle, 2023), effsize (Torchiano, 2020), car (Fox et al., 2024), and Rmisc (Warnes, 2016). Finally, the formatting and reporting of results were generated using flextable (Gohel, 2024) and officer (Gohel, 2024).

The initial code for data analysis was developed independently by the authors. Additional calculations performed during various stages of manuscript revision were conducted with the consultative assistance of the Qwen 3.7-Plus large language model[1]. All statistical decisions, interpretation of the results, and formulation of the conclusions were made independently by the authors. Furthermore, the aforementioned language model was utilized to translate specific sections of the manuscript into English, as the authors are non-native English speakers.

[1] Alibaba Cloud (2026). Qwen3.7-Plus [Large language model]. https://qwen.ai/

Results

Tables 2–4 present the results of analyzing differences in the expression of various values within the Shwartz model. In particular, the differences between groups of respondents with low and high levels of aggression, including physical aggression (*Table 2*), anger (*Table 3*), and hostility (*Table 4*), are examined.

Table 2

Robustness of Group Differences in Schwartz's Value Dimensions: Proportion of Statistically Significant Results (p < .05) Across 100 Bootstrap Samples, Comparing Respondents with High vs. Low Physical Aggression

Values (PVQ-RR)	*t*	Mean Diff [95% CI]	*U (Bst)* (%)	*t (Bst)* (%)	Cohen's *d* *M* [95% CI]
Self-Direction — Thought	–1.01	–.21[–.62, .2]	0	5	–.09 [–.10, –.08]
Self-Direction — Action	.474	.1 [–.31, .5]	0	0	.04 [.03, .05]
Stimulation	4.28***	.97[.52, 1.42]	100	100	.39 [.37, .40]
Hedonism	2.10*	.44[.03, .85]	0	0	.19 [.18, .20]
Achievement	.41	.09[–.34, .52]	0	0	.04 [.03, .05]
Power — Dominance	9.74***	2.95[2.35, 3.54]	100	100	.89 [.88, .90]
Power — Resources	4.59***	1.10[.63, 1.57]	100	100	.41 [.40, .42]
Face	–.08	–.02[–.51, .47]	0	0	–.01 [–.03, .00]
Security — Personal	–6.80***	–1.55[–2.00, –1.10]	100	100	–.64 [–.65, –.63]
Security — Societal	–3.75***	–.93[–1.41, –.44]	97	98	–.36 [–.37, –.35]
Tradition	–2.40*	–.75[–1.36, –.13]	0	1	–.23 [–.24, –.21]
Conformity — Rules	–7.82***	–2.16[–2.71, –1.61]	100	100	–.74 [–.75, –.73]
Conformity — Interpersonal	–5.59***	–1.52[–2.06, –.98]	100	100	–.53 [–.54, –.52]
Humility	–5.55***	–1.45[–1.96, –.93]	100	100	–.53 [–.54, –.52]
Universalism — Concern	–4.33***	–1.01[–1.47, –.55]	100	100	–.40 [–.42, –.39]
Universalism — Nature	–4.98***	–1.23[–1.72, –.74]	100	100	–.45 [–.46, –.44]
Universalism — Tolerance	–7.97***	–2.11[–2.64, –1.59]	100	100	–.78 [–.80, –.77]
Benevolence — Caring	–5.35***	–1.24[–1.69, –.78]	100	100	–.52 [–.53, –.50]
Benevolence — Dependability	–4.38***	–.97[–1.41, –.53]	100	100	–.42 [–.43, –.41]

*Note. * — p<.05; ** — p<.01; *** — p < .001; Mean Diff = Mean difference; Bst = Bootstrap; t and Mean Diff [95%CI] are calculated on a full sample. U (Bst) and t (Bst) are calculated on weighted samples. Proportions indicate the percentage of 100 bootstrap replications in which group differences reached statistical significance (p < .05, two-tailed). Effect sizes (Cohen's d) represent the mean d and its 95CI across 100 replications.*

Table 3

Robustness of Group Differences in Schwartz's Value Dimensions: Proportion of Statistically Significant Results (p < 0.05) Across 100 Bootstrap Samples, Comparing Respondents with High vs. Low Anger

Values (PVQ-RR)	*t*	Mean Diff [95% CI]	*U (Bst)* (%)	*t (Bst)* (%)	Cohen's *d* *M* [95% CI]
Self-Direction — Thought	2.68**	.44[.12, .76]	84	38	.23 [.22, .24]
Self-Direction — Action	4.40***	.78[.43, 1.13]	100	100	.36 [.35, .37]
Stimulation	4.37***	.90[.49, 1.30]	100	100	.37 [.36, .38]
Hedonism	5.27***	1.00[.63, 1.38]	100	100	.44 [.42, .45]
Achievement	5.84***	1.18[.78, 1.57]	100	100	.48 [.47, .49]
Power — Dominance	10.26***	2.95[2.38, 3.52]	100	100	.88 [.87, .89]
Power — Resources	7.71***	1.81[1.35, 2.28]	100	100	.62 [.61, .64]
Face	4.86***	1.15[.68, 1.61]	100	100	.40 [.39, .41]
Security — Personal	–3.04**	–.64[–1.06, –.23]	75	88	–.27[–.28, –.25]
Security — Societal	–1.18	–.27[–.72,.18]	0	0	–.09[–.10, –.08]
Tradition	–3.70***	–1.13[–1.74,–.53]	99	97	–.31[–.32, –.30]
Conformity — Rules	–5.73***	–1.55[–2.08,–1.01]	100	100	–.51[–.52, –.50]
Conformity — Interpersonal	–5.07***	–1.51[–2.09, –.92]	100	100	–.46[–.47, –.45]
Humility	–7.87***	–2.02[–2.53, –1.51]	100	100	–.70[–.71, –.69]
Universalism — Concern	–.11	–.02[–.43, .38]	0	0	–.02[–.03, –.01]
Universalism — Nature	–4.13***	–1.02[–1.51, –.53]	100	100	–.35[–.36, –.33]
Universalism — Tolerance	–4.18***	–1.01[–1.49, –.53]	95	99	–.36[–.37, –.35]
Benevolence — Caring	.43	.08[–.27, .42]	0	0	.03[.02, .04]
Benevolence — Dependability	3.60***	.63[.29, .98]	100	90	.28[.27, .30]

*Note. * — p<.05; ** — p<.01; *** — p < .001; Mean Diff = Mean difference; Bst = Bootstrap; t and Mean Diff [95%CI] are calculated on a full sample. U (Bst) and t (Bst) are calculated on weighted samples. Proportions indicate the percentage of 100 bootstrap replications in which group differences reached statistical significance (p < .05, two-tailed). Effect sizes (Cohen's d) represent the mean d and its 95CI across 100 replications.*

According to the results of the analysis, it was established that the expression of different values within the sample of Russian university students with high and low physical aggression has three different distribution patterns. Respondents with high physical aggression had significantly higher levels of Stimulation, Power (dominance and resources). Respondents with low physical aggression had significantly greater values of Security, Conformity, Humility, and Universalism.

Table 4

Robustness of Group Differences in Schwartz's Value Dimensions: Proportion of Statistically Significant Results (p < 0.05) Across 100 Bootstrap Samples, Comparing Respondents with High vs. Low Hostility

Values (PVQ-RR)	*t*	Mean Diff [95% CI]	*U (Bts)* *(%)*	*t (Bts)* *(%)*	Cohen's *d* M [95% CI]
Self-Direction — Thought	1.44	.23[–.08, .54]	0	0	.10 [.09, .11]
Self-Direction — Action	.19	.03[–.31, .38]	0	0	.02 [.01, .03]
Stimulation	3.48***	.67 [.29, 1.05]	94	87	.27 [.25, .28]
Hedonism	3.46***	.63 [.27, .99]	83	95	.28 [.27, .29]
Achievement	3.71***	.70 [.33, 1.08]	97	96	.30 [.29, .31]
Power — Dominance	8.61***	2.29 [1.77, 2.82]	100	100	.71 [.70, .72]
Power — Resources	6.12***	1.34 [0.91, 1.78]	100	100	.49 [.48, .51]
Face	9.03***	1.81 [1.42, 2.20]	100	100	.69 [.68, .70]
Security — Personal	–1.36	–.25[–.62, .11]	0	0	–.11[–.12, –.09]
Security — Societal	–.05	–.01[–.42, .40]	0	0	–.01 [–.02, .00]
Tradition	–2.28*	–.67[–1.25, –.09]	0	6	–.19 –.20, –.18]
Conformity — Rules	–3.10**	–.73[–1.20, –.27]	71	68	–.25 [–.26, –.24]
Conformity — Interpersonal	2.33*	.58[.09, 1.07]	19	0	.18 [.17, .19]
Humility	–1.04	–.25[–.71, .22]	0	0	–.09 [–.10, –.07]
Universalism — Concern	1.12	.20[–.15, .56]	0	0	.09 [.08, .10]
Universalism — Nature	–2.25*	–.46[–.86, –.06]	46	42	–.18 [–.19, –.16]
Universalism — Tolerance	–3.68***	–.75[–1.15, –.35]	98	94	–.31 [–.32, –.29]
Benevolence — Caring	–1.09	–.20[–.55, .16]	0	0	–.09 [–.10, –.08]
Benevolence — Dependability	.09	.02[–.33, .36]	0	0	.01 [.00, .02]

*Note. * — p<.05; ** — p<.01; *** — p < .001; Mean Diff = Mean difference; Bst = Bootstrap; t and Mean Diff [95%CI] are calculated on a full sample. U (Bst) and t (Bst) are calculated on weighted samples. Proportions indicate the percentage of 100 bootstrap replications in which group differences reached statistical significance (p < .05, two-tailed). Effect sizes (Cohen's d) represent the mean d and its 95CI across 100 replications.*

The expression of different values in the context of various levels of anger also shows heterogeneity: respondents with a high level of anger are characterized by the expression of value orientations such as Self-Direction — Action, Stimulation, Hedonism, Achievement, Power (dominance and resources), and Face. A low level of anger, on the other hand, is associated with the expression of values like Tradition,

Conformity (rules and interpersonal), Humility, Universalism (nature and tolerance), and Benevolence — Dependability.

The respondents characterized by a high level of hostility revealed preferences for the following values: Power (dominance and resources) and Face. The values of Stimulation, Hedonism, and Achievement can also be attributed to this group of results. However, the pattern is established in slightly fewer cases, and the mean effect size ranged from .27–.30. Low hostility corresponds with the expression of Universalism — Tolerance value.

Discussion

The results of this study are generally consistent with previous research on the relationship between values and aggression. Specifically, we found that as levels of indicators of aggression (physical aggression, anger, and hostility) increase, so does the importance of values related to Stimulation and Power (dominance and resources). These findings are in line with previous studies by McDonald et al. (2015), Rubel-Lifschitz et al. (2021), and Seddig and Davidov (2018).

Stimulation as a value is often associated with aggression in various forms, as supported by the results by Schwartz (2022) and Bardi and Schwartz (2003). In other words, a desire for excitement, novelty, and risk can lead to an increased likelihood of aggressive behavior. This may be because individuals with a high need for Stimulation seek intense experiences, and aggression (particularly physical aggression) may seem like a way to obtain these experiences. Additionally, impulsivity, which often goes hand in hand with Stimulation, can contribute to rapid and uncontrolled aggression, further supporting the idea that Stimulation is linked to aggression. Impulsivity has also been linked to deviance and aggression, as suggested by studies by Rean et al. (2016) and Bresin (2019).

Power (dominance, resources) can be expressed through various forms of aggression. This is supported by global evidence that power is a significant factor for aggressive individuals (Knafo et al., 2008; Rokas et al., Sagiv et al., 2022). In this context, aggression can serve to establish or maintain a dominant position, particularly when power is perceived as a scarce resource to be fought for.

The general trend is that people who prioritize "personal focus" were more likely to exhibit higher levels of aggression. In contrast, those who prioritize "social focus" tended to have lower rates of aggression. Given that the study measured various indicators of aggression, it is important to note that differences were only found in some of these indicators, not all of them.

The association between Openness to Change values and aggression indicators merits attention. Notably, differences in Self-Direction emerged solely in relation to anger. Given that Self-Direction reflects "freedom to choose one's actions" (Schwartz et al., 2012), it may be inferred that anger functions as a guide for goal pursuit within socially acceptable boundaries. This is evidenced by the lack of significant associations with physical aggression measures. These results align indirectly with findings on anger's role in goal-setting compared to sadness, particularly regarding elevated aspirations (Chong & Park, 2017).

The values of Hedonism, Achievement, and Face, as shown by our results, are linked to the parameters of anger and hostility, rather than physical aggression. This suggests that physical aggression may not be used to achieve goals, whereas other forms, such as anger and hostility, play an important role. It is possible that individuals characterized by these values may resort to dominant behavior in order to gain a competitive advantage in social interactions. At the same time, we can assume the strategic nature of their aggressive behavior, where direct confrontation may not be the optimal way to achieve goals. Instead, anger and hostility can be used to establish status and influence. This aligns with the hypothesis within evolutionary and ecological approaches regarding the adaptive role of aggressive behavior as a high-risk strategy in social interactions (Pyatnitsky & Enikolopov, 2024).

The relationship between values from the Self-Transcendence and Conservation domains and aggression indicators warrants examination. According to the results of this study, Self-Transcendence values (Universalism and Benevolence) were negatively associated with physical aggression, with the notable exception of Tradition—a Conservation value — which did not demonstrate a significant association. In the context of varying levels of hostility, significant differences were identified exclusively for the Universalism–Tolerance parameter. Regarding anger, no significant differences were observed in values such as Security–Societal (Conservation domain), Universalism–Concern, and Benevolence–Caring (both Self-Transcendence domain).

The greater expression of Self-Transcendence values among respondents with low physical aggression is consistent with findings from previous research on Russian samples (Meshkova, 2018). However, it is important to acknowledge that this pattern does not rule out potential bias, as evidenced by the absence of significant differences in hostility across nearly all measured values. Data concerning the emotional component of aggression (anger) generally confirm the inverse relationship between Self-Transcendence values and aggression, though several exceptions were identified. Notably, the non-significant differences in Security–Societal (Conservation), Universalism–Concern, and Benevolence–Caring (Self-Transcendence) values across anger levels suggest that these values do not preclude the manifestation of anger in defense of group interests.

An exception to the trends discussed above are the findings obtained from measuring the value of Benevolence–Dependability, which refers to an individual's desire to be a reliable and trustworthy member of a group. The high expression of this value orientation is associated with low physical aggression and high anger. This suggests that students with a strong orientation towards this value may be more likely to favor their own group, potentially leading to in-group bias (Wang et al., 2016).

It has been established that the relationship between values and aggression depends on the cultural context. In societies with a strong focus on social control and collectivism, values such as Security play a more significant role in deterring aggression compared to cultures that promote individualism and competition (Bergeron & Schneider, 2005; Bergmüller, 2013; Ioanna & Kokkinos, 2024). This difference can be attributed to the correlation between individual value orientations and the values of a reference group. According to Radchikova and Sorokova (2023), when a person's values align with those of their reference group, it contributes to their subjective well-

being and self-esteem, as also demonstrated by Rean et al. (2024). Considering that the participants in our study were university students in the Russian Federation, it can reasonably be argued that the established relationship between social focus values and a decrease in aggression indicators indicates the consistency of individual value attitudes among students and the need for social stability and justice in modern Russian society (VCIOM, 2024).

Conclusion

Within the framework of the study, conducted on a sample of Russian student, differences in the expression of values according to Schwartz's value theory were identified in the context of indicators of aggression. It was found that as the level of each indicator of aggression (physical aggression, anger, and hostility) increases, the importance of values such as Stimulation and Power (dominance and resources) also increases, while the importance of Universalism–Tolerance decreases.

Furthermore, it was discovered that a preference for personal-focused values is associated with higher levels of aggression, while a preference for social-focused values is linked to lower levels. Moreover, it has been demonstrated that social-focused values are negatively related to physical aggression, except for the value of Tradition, which connection was not robust. In the context of hostility, differences are only established by the Universalism–Tolerance parameter. In groups with varying degrees of anger, there were no significant differences in values such as Security–Societal, Universalism–Concern, and Benevolence–Caring.

Limitations

The limitations of this study include a number of points. First, we should note the limitations associated with the use of self-reporting techniques: respondents may give a more socially acceptable answer, rather than being truthful, or subjects may not be able to assess themselves accurately. Second, the study presents a sample that is biased towards a certain gender. Despite the fact that the sample may be considered representative of students with social and humanitarian educational backgrounds, the results may not be generalizable to the relationship between aggression and values among all students, but mainly among young women. Thirdly, it is worth noting the low reliability of several scales (five subscales ranged from .60 to .70, and .50 to .60 for two scales) in a specific sample, which requires further verification of the results obtained in other samples. Fourth, the analysis methodology did not consider side variables, which is also a limitation. We also note that resampling techniques cannot estimate the probability of events that are not present in the original sample, limiting their applicability for analyzing rare or extreme events, which is another limitation. Although the overlap between high aggression groups is relatively low (31–36%), the moderate inter-scale correlations ($r = .41–.58$) indicate that the subscales are related. The differential patterns observed across subscales should therefore be interpreted as complementary facets of a broader aggression profile. Future studies employing multivariate approaches (*e.g.*, MANOVA or structural equation modeling) could further disentangle the unique contributions of each ag-

gression facet. The cross-sectional design of this study precludes causal inferences regarding the relationships between personal values and aggression; longitudinal research is needed to establish temporal precedence and directionality. These shortcomings could be addressed in future studies by adjusting for gender, age, and using covariate-adjusted models.

Informed Consent

Respondents confirmed their consent to participate in the study in the online survey.

Conflict of Interest

The authors declare no conflict of interest

Ethics Statement

The study obtained ethics approval of the Ethics Committee of the Federal Scientific Center for Psychological and Interdisciplinary Research (protocol number 9 of 31.01.2024).

Author Contributions

I.A.K. and A.A.R. conceived of the idea and developed the methodology. A.A.R. supervised the research and organized the collection of empirical data. I.A.K. performed the data analysis and drafted the Methods, Results, and Discussion sections, while A.L.L. drafted the Introduction. I.A.K. reviewed and edited the text, and prepared the final version of the manuscript. All authors discussed the results and contributed to the final manuscript.

References

Baeva, I.A., Ulyanina, O.A., Vikhristiuk, O.V., Gaiazova, L.A., Radchikova, N.P., & Faizullina, K.A. (2022). Faktory sotsial'no-psikhologicheskoi dezadaptatsii i tsennostnye orientatsii podrostkov i molodezhi [Factors of socio-psychological maladaptation and value orientations of adolescents and youth]. *Izvestiia RGPU im. A.I. Gertsena* [Izvestia: Herzen University Journal of Humanities & Sciences], *(206)*, 228–242.

Bardi, A., & Schwartz, S.H. (2003). Values and behavior: Strength and structure of relations. *Personality and Social Psychology Bulletin, 29*(10), 1207–1220. https://doi.org/10.1177/0146167203254602

Baron, R.A., & Richardson, D.R. (1994). *Human aggression.* Springer Science & Business Media.

Benish-Weisman, M. (2015). The interplay between values and aggression in adolescence: A longitudinal study. *Developmental Psychology, 51*(5), 677–687. https://doi.org/10.1037/a0038921

Benish-Weisman, M., & McDonald, K.L. (2015). Private self-consciousness and gender moderate how adolescents' values relate to aggression. *Social Development, 24*(4), 766–781. https://doi.org/10.1111/sode.12120

Benish-Weisman, M., Daniel, E., & Knafo-Noam, A. (2017). The relations between values and aggression: A developmental perspective. In T.-V. Luijkx, L. Cieciuch, & M. Benish-Weisman (Eds.), *Values and behavior: Taking a cross cultural perspective* (pp. 97–114). Springer. https://doi.org/10.1007/978-3-319-56352-7_6

Bergeron, N., & Schneider, B.H. (2005). Explaining cross-national differences in peer-directed aggression: A quantitative synthesis. *Aggressive Behavior, 31*, 116–137. https://doi.org/10.1002/ab.20049

Bergmueller, S. (2013). The relationship between cultural individualism–collectivism and student aggression across 62 countries. *Aggressive Behavior, 39*(3), 182–200. https://doi.org/10.1002/ab.21474

Bresin, K. (2019). Impulsivity and aggression: A meta-analysis using the UPPS model of impulsivity. *Aggression and Violent Behavior, 48*, 124–140. https://doi.org/10.1016/j.avb.2019.08.003

Buss, A.H. (1961). *The psychology of aggression.* Wiley.

Buss, A.H., & Perry, M. (1992). The aggression questionnaire. *Journal of Personality and Social Psychology, 63*(3), 452–459. https://doi.org/10.1037/0022-3514.63.3.452

Chong, S.H., & Park, G. (2017). The differential effects of incidental anger and sadness on goal regulation. *Learning and Motivation, 58*, 1–15. https://doi.org/10.1016/j.lmot.2017.04.002

Chudova, N.V. (2014). Issledovanie tsennostnykh osnovanii agressivnosti [Study of value foundations of aggressiveness]. *Prikladnaia iuridicheskaia psikhologiia*[Applied Legal Psychology], *(2)*, 48–59.

Cohen, J. (2013). *Statistical power analysis for the behavioral sciences.* Routledge. https://doi.org/10.4324/9780203771587

Efron, B., & Tibshirani, R.J. (1994). *An introduction to the bootstrap*. Chapman and Hall/CRC. https://doi.org/10.1201/9780429246593

Enikolopov, S.N., & Tsibulsky, N.P. (2007). Psikhometricheskii analiz russkoyazychnoy versii Oprosnika diagnostiki agressii A. Bassa i M. Perri [Psychometric analysis of the Russian version of the Buss-Perry Aggression Questionnaire]. *Psikhologicheskii zhurnal* [Psychological journal], *28*(1), 115–124.

Fox, J., Weisberg, S., & Price, B. (2024). *car: Companion to applied regression* (R package version 3.1-2) [Computer software]. https://CRAN.R-project.org/package=car

Gohel, D. (2024). *flextable: Functions for tabular reporting* (R package version 0.9.6) [Computer software]. https://CRAN.R-project.org/package=flextable

Gohel, D. (2024). *officer: Manipulation of Microsoft Word and PowerPoint documents* (R package version 0.6.6) [Computer software]. https://CRAN.R-project.org/package=officer

Ioanna, V., & Kokkinos, C.M. (2024). Relational aggression in Greek university students: A latent profile analysis of students' cultural orientations. *European Review of Applied Psychology, 74*(4), 100931. https://doi.org/10.1016/j.erap.2024.100931

Karandashev, V.N. (2004). *Metodika Shvartsa dlia izucheniia tsennostei lichnosti: kontseptsiia i metodicheskoe rukovodstvo* [Schwartz's method for studying personal values: Concept and methodological guide]. Rech.

Khusainova, S.V., & Khakimzyanov, R.N. (2019). Issledovanie psikhologicheskikh osobennostei lichnosti sklonnoi k deviantnomu povedeniiu [Study of personality traits prone to deviant behavior]. *Kazanskii pedagogicheskii zhurnal* [Kazan Pedagogical Journal], *6*(137), 195–201.

Knafo, A., Daniel, E., & Khoury-Kassabri, M. (2008). Values as protective factors against violent behavior in Jewish and Arab high schools in Israel. *Child Development, 79*(3), 652–667. https://doi.org/10.1111/j.1467-8624.2008.01149.x

Korneev, A.A., & Krichevets, A.N. (2011). Usloviia primenimosti kriteriev St'iudenta i Manna-Uitni [Applicability conditions of Student's and Mann–Whitney criteria]. *Psikhologicheskii zhurnal* [Psychological journal], 32(1), 97–110.

Krahe, B. (2013). *The social psychology of aggression*. Taylor & Francis.

Ludecke, D. (2023). *sjmisc: Data and variable transformation functions* (R package version 2.8.9) [Computer software]. https://CRAN.R-project.org/package=sjmisc

McDonald, K.L., Lochman, J.E., Boxmeyer, C.L., Powell, N., & Sallee, M. (2015). The social values of aggressive–prosocial youth. *Journal of Youth and Adolescence, 44*, 2245–2256. https://doi.org/10.1007/s10964-015-0344-7

Meshkova, N.V. (2018). Osobennosti vzaimosvyazi antisosial'no napravlennoi kreativnosti i tsennostei u podrostkov s raznym urovnem agressii [Features of the relationship between antisocial creativity and values in adolescents with different levels of aggression]. *Psikhologo-pedagogicheskie*

issledovaniia [Psychological-pedagogical research], *10*(2), 77–87. https://doi.org/10.17759/psyedu.2018100207

Myburgh, C., Poggenpoel, M., & Fourie, C.M. (2020). Predictors of aggression of university students. *Health SA Gesondheid, 25*, 1376. https://doi.org/10.4102/hsag.v25i0.1376

Obukhovskaya, D.I., Konopleva, I.N. (2023). Aktsentuatsii kharaktera, tsennostnye orientatsii i agressivnost' podrostkov kak faktory riska sovershenniia pravonarushenii nesovershennoletnimi [The Character Accentuations, Value Orientations and Agressiveness of Adolescents as the Risk Factors of the Committing of Offenses by Juveniles]. *Psikhologiia i pravo* [Psychology and Law], *13*(4), 115–131. https://doi.org/10.17759/psylaw.2023130409

Oshchepkov, A.A., Piskunova, E.N., & Salakhova, V.B. (2023). Osobennosti vzaimosviazi zhiznesposobnosti i tsennostnykh orientatsii lichnosti molodezhi, sklonnoi k deviantnomu povedeniiu [Resilience and value orientations in deviant youth]. *Alma Mater*, *(11)*, 31–38. https://doi.org/10.20339/AM.11-23.031

Posit Team. (2024). *RStudio: Integrated Development Environment for R*. Posit Software, PBC. http://www.posit.co/

Pyatnitsky, N.Yu., & Enikolopov, S.N. (2024). Antisosial'noe rasstroistvo lichnosti kak forma adaptatsionnoi strategii [Antisocial personality disorder as an adaptation strategy]. *Psikhiatriia* [Psychiatry], *22*(3), 63–72.

R Core Team. (2023). *R: A language and environment for statistical computing*. R Foundation for Statistical Computing. https://www.R-project.org/

Radchikova, N.P., & Sorokova, M.G. (2023). Tsennostnye orientatsii v psikhologicheskom blagopoluchii podrostka: itogi strukturnogo modelirovaniia [Value orientations in adolescent psychological well-being: Structural modeling results]. *Sotsial'nye nauki i detstvo* [Social sciences and childhood], *4*(1), 7–17. https://doi.org/10.17759/ssc.2023040101

Rean, A.A. (Ed.). (2016). *Psikhologiia deviantnosti. Deti. Obshchestvo. Zakon: monografiia* [Psychology of deviance. Children. Society. Law: A monograph]. IUNITI-DANA.

Rean, A. A., Egorova, A.V., Konovalov, I.A., Kuz'min, R.G., Schevchenko, A.O., Stavtsev, A.A., & Rean, K.A. (2024). Vzaimosviaz' samootsenki (samouvazheniia) i tsennostnykh orientatsii rossiiskikh studentov [The relationship between self-esteem (self-respect) and value orientations of Russian students]. *Integratsiia obrazovaniia* [Integration of education], *28*(4), 498–513. https://doi.org/10.15507/1991-9468.129.028.202404.498-513

Revelle, W. (2023). *psych: Procedures for psychological, psychometric, and personality research* (Version 2.3.6) [R package]. Northwestern University. https://CRAN.R-project.org/package=psych

Rubel-Lifschitz, T., Benish-Weisman, M., Torres, C. V., & McDonald, K. (2021). The revealing effect of power: Popularity moderates the associations of personal values with aggression in adolescence. *Journal of Personality, 89*(4), 786–802. https://doi.org/10.1111/jopy.12615

Sagiv, L., & Schwartz, S.H. (2022). Personal values across cultures. *Annual Review of Psychology, 73*(1), 517–546. https://doi.org/10.1146/annurev-psych-020821-121934

Salakhova, V.B., Garanyan, N.G., Grishina, N.V., & Aslanova, E.R. (2016). Features of social attitudes and value orientations of youths and adolescents prone to auto-aggressive behavior. *International Journal of Environmental and Science Education, 11*(16), 9017–9025.

Schwartz, S.H. (1992). Universals in the content and structure of values: Theoretical advances and empirical tests in 20 countries. In M. P. Zanna (Ed.), *Advances in experimental social psychology* (pp. 1–65). Academic Press. https://doi.org/10.1016/S0065-2601(08)60281-6

Schwartz, S.H. (2006). Basic Human Values: An Overview.

Schwartz, S.H. (2012). An overview of the Schwartz theory of basic values. *Online Readings in Psychology and Culture, 2*(1), 11. https://doi.org/10.9707/2307-0919.1116

Schwartz, S.H. (2022). Utochnennaia teoriia bazovykh tsennostei. In S. Rokkas & L. Sagiv (Eds.), *Tsennosti i povedenie: Kross-kul'turnyi podkhod* (pp. 71–100). Institut psikhologii RAN.

Schwartz, S.H., Butenko, T.P., Sedova, D.S., & Lipatova, A.S. (2012). Utochnennaia teoriia bazovykh individual'nykh tsennostei: primenenie v Rossii [Refined theory of basic individual values: Applica-

tion in Russia]. *Psikhologiia. Zhurnal vysshei shkoly ekonomiki* [Psychology. Journal of the Higher School of Economics], *9*(2), 43–70.

Seddig, D., & Davidov, E. (2018). Values, attitudes toward interpersonal violence, and interpersonal violent behavior. *Frontiers in Psychology, 9*, 604. https://doi.org/10.3389/fpsyg.2018.00604

Sobkin, V.S., & Kalashnikova, E.A. (2021). K voprosu o mezhpokolencheskikh razlichiiakh v sotsial'nom samochuvstvii podrostkov: zhiznennaia pozitsiia i otsenka znachimosti ugroz. [On the question of intergenerational differences in the social well-being of adolescents: Life position and assessment of threat significance]. *Psikhologicheskaia nauka i obrazovanie* [Psychological Science and Education], *26*(3), 54–69. https://doi.org/10.17759/pse.2021260303

Torchiano, M. (2020). *effsize: Efficient effect size computation* (Version 0.8.1) [R package]. https://doi.org/10.5281/zenodo.1480624

Usacheva, I.V., & Parshutin, I.A. (2023). Tsennostnye orientatsii kursantov MVD Rossii kak regulator lichnostnoi agressivnosti [Value orientations of cadets as a regulator of aggressiveness]. *Rossiyskii deviantologicheskii zhurnal* [Russian deviantological journal], *3*(3), 281–289.

VCIOM (2024). Sotsial'naia spravedlivost': monitoring [Social justice: Monitoring]. https://wciom.ru/analytical-reviews/analiticheskii-obzor/socialnaja-spravedlivost-monitoring

Velensia, M., Soetikno, N., & Chandhika, J. (2021). The role of conformity towards the self-control on adolescence cyberbullying. In *International Conference on Economics, Business, Social, and Humanities (ICEBSH 2021)* (pp. 281–286). Atlantis Press. https://doi.org/10.2991/assehr.k.210805.045

Wang, L., Zhu, C., Zhang, L., Zhang, S., & Zhang, Y. (2016). Ingroup favoritism or the black sheep effect: Perceived intentions modulate subjective responses to aggressive interactions. *Neuroscience Research, 108*, 46–54. https://doi.org/10.1016/j.neures.2016.01.006

Warnes, G.R. (2016). *Rmisc: Ryan miscellaneous* (R package version 1.0.1) [Computer software]. https://CRAN.R-project.org/package=Rmisc

Wickham, H., & Henry, L. (2023). *purrr: A functional programming toolkit* (R package version 1.0.2) [Computer software]. https://CRAN.R-project.org/package=purrr

Wickham, H., & Bryan, J. (2023). *readxl: Read Excel files* (Version 1.4.2) [R package]. https://CRAN.R-project.org/package=readxl

Wickham, H., François, R., Henry, L., Müller, K., & Vaughan, D. (2023). *dplyr: A grammar of data manipulation* (Version 1.1.2) [R package]. https://CRAN.R-project.org/package=dplyr

Youn, H., & Lee, Y.H. (2016). Values of people with primary and secondary psychopathic tendency: Focused on Schwartz's value theory. *Korean Journal of Clinical Psychology, 35*(2), 473–497. https://doi.org/10.15842/kjcp.2016.35.2.008

Original manuscript received May 16, 2025
Revised manuscript accepted September 14, 2026
First published online September 15, 2026

To cite this article: Konovalov, I.A., Linkov, A.L., Rean, A.A. (2026). The Values System of Russian University Students Compared with Their Level of Aggression. *Psychology in Russia: State of the Art*, *19*(3), 131–150. DOI: 10.11621/pir.2026.0307

Appendix A. Results of Overlap Analysis

Table 1

Overlap Between Physical and Anger Groups (N = 2,328)

Physical	High anger	Mid anger	Low anger	Total
High	50 (31.4%)	102 (64.2%)	7 (4.4%)	159
Mid	100 (8.6%)	747 (64.5%)	312 (26.9%)	1159
Low	34 (3.4%)	413 (40.9%)	563 (55.7%)	1010

Note. Values represent n (%). Percentages are row percentages. N = 2,328.

Table 2

Overlap Between Physical and Hostility Groups (N = 2,328)

Physical	High hostility	Mid hostility	Low hostility	Total
High	52 (32.7%)	95 (59.7%)	12 (7.5%)	159
Mid	100 (8.6%)	769 (66.4%)	290 (25%)	1159
Low	49 (4.9%)	456 (45.1%)	505 (50%)	1010

Note. Values represent n (%). Percentages are row percentages. N = 2,328.

Table 3

Overlap Between Anger and Hostility Groups (N = 2,328)

Anger	High hostility	Mid hostility	Low hostility	Total
High	66 (35.9%)	104 (56.5%)	14 (7.6%)	184
Mid	123 (9.7%)	880 (69.7%)	259 (20.5%)	1262
Low	12 (1.4%)	336 (38.1%)	534 (60.5%)	882

Note. Values represent n (%). Percentages are row percentages. N = 2,328.

Table 4

Summary of Overlap Between Aggression Subscale Groups (N = 2,328)

Subscale pair	*n* High (row)	High–High overlap	High–Mid	High–Low	*r*
Physical × Anger	159	50 (31.4%)	102 (64.2%)	7 (4.4%)	.46***
Physical × Hostility	159	52 (32.7%)	95 (59.7%)	12 (7.5%)	.41***
Anger × Hostility	184	66 (35.9%)	104 (56.5%)	14 (7.6%)	.58***

*Note. High–High overlap = percentage of respondents in the High group of the first subscale who were also High on the second subscale. r = Pearson correlation between continuous subscale scores. ***p >.001.*