ChinaRxiv

AI translation · View original & related papers at
chinarxiv.org/items/chinaxiv-202305.00016

Measurement and Impact of Academic Involution Among University Students

Authors: Liu Yingying, Xie Xintong, Zhang Shunmin, Zhang Shunmin

Date: 2023-05-04T00:00:00+00:00

Abstract

Although academic involution may potentially undermine university talent cultivation, reliable measurement instruments remain lacking for its empirical examination. This study developed a questionnaire for assessing academic involution among university students and validated its psychometric properties through three studies. Study 1 constructed an initial item pool through literature analysis, real-world observation, and interviews, and subsequently screened items based on pilot testing data from 338 university students. Study 2 established a final 16-item, three-factor questionnaire (sacrificing intrinsic values for external gains, competitive disqualification, and detachment from practical reality) via exploratory factor analysis of a large sample (N = 3000) and confirmatory factor analysis of an independent sample (N = 571). Results from the 3000-participant sample indicated that the majority of students exhibited high levels of academic involution. Individuals with high academic involution manifested stronger zero-sum beliefs, elevated trait anxiety, diminished life satisfaction, and poorer sleep quality, yet did not demonstrate enhanced creativity tendencies. Study 3, utilizing data from 99 participants, revealed that the questionnaire's test-retest reliability coefficient was 0.83. This investigation not only developed a reliable measurement instrument but also demonstrated that academic involution is prevalent among university students and exerts significant detrimental effects.

Full Text

The Measurement and Impact of Academic Involution Among College Students

Liu Yingying, Xie Xintong, Zhang Shunmin
(Department of Psychology and Behavioral Sciences, Zhejiang University, Hangzhou 310058)

Abstract

Although academic involution may harm the cultivation of talent in higher education, reliable measurement tools remain unavailable. This paper developed a College Students' Academic Involution Scale (CAIS) and examined its reliability and validity through three studies. Study 1 constructed an initial item pool based on literature analysis, real-world observations, and interviews, and screened items using data from 338 undergraduates. Study 2 established a final 16-item, three-factor questionnaire (unwilling hardworking, excessive competition, and surface learning) through exploratory factor analysis with a large sample (N = 3,000) and confirmatory factor analysis with an independent sample (N = 571). The 3,000-participant sample revealed that most students exhibited high levels of academic involution. Specifically, individuals with high CAIS scores showed stronger zero-sum beliefs, higher trait anxiety, lower life satisfaction, and poorer sleep quality, yet did not demonstrate higher creative potential. Study 3 found a test-retest reliability of 0.83 based on data from 99 participants. This research not only provides a reliable measurement instrument but also reveals that academic involution is both prevalent among college students and demonstrably harmful.

Keywords: academic involution, scale development, trait anxiety, sleep, creativity

---

1. Introduction

Involution, as a term describing "excessive competition" and "devaluation of effort," has attracted widespread and sustained social attention in recent years. A photograph that circulated in 2020 sparked heated discussion about involution, depicting a college student working on a laptop while riding a bicycle. This startling image prompted reflection: Has academic competition in universities become so intense? As the term gained popularity, deeper questions emerged. Does academic involution truly exist? How pervasive is it among university students? What impact does it have on talent development? What are its underlying causes? Addressing these questions requires a clear definition of involution and effective measurement tools.

The concept of involution was introduced to the social sciences in 1936 by anthropologist Goldenweiser to describe cultural patterns that, upon reaching a definitive form, fail to evolve into new patterns but instead continue developing only toward greater internal complexity. Geertz (1963) subsequently applied the term to agriculture in Java, where countless laborers concentrated on limited rice production. Although per-unit yields increased, living standards improved only marginally. Huang (2002) was the first scholar to introduce the concept to Chinese research, annotating involution as occurring when the marginal product of labor begins to decrease. Despite slight differences from Geertz's definition (Liu and Qiu, 2004), Huang's interpretation shares common ground with

Geertz's. Huang (2021) emphasized four characteristics of involution: (1) increasingly high labor input, (2) decreasing marginal returns, (3) formation of a rigid closed system, and (4) resistance to or exclusion of qualitative change.

While most scholars view involution as having negative consequences, a few have attempted positive interpretations (Guo, 2007), describing Geertz's concept as a "process of self-overcoming." However, this positive reading misrepresents Geertz's original meaning. The phrase "self-defeating process" was incorrectly translated as "self-overcoming." Self-defeating means "counterproductive" or "self-thwarting," not "self-overcoming." In Geertz's description, Javanese agricultural producers were helpless and compelled, pouring more labor into limited land while barely surviving and seeing no hope for change (McCullough, 2019).

In summary, involution's basic features are: (1) constraints of an established framework, (2) continuously escalating input, and (3) lack of substantive gains or development. Goldenweiser's "definitive form of cultural patterns," Geertz's "limited rice production," and Huang's (2021) "closed system" all constitute established frameworks. Meanwhile, Goldenweiser's "becoming more complex internally," Geertz's "concentration of countless laborers," and Huang's (2021) "increasingly high labor input" all reflect escalating input. From an individual perspective, involution manifests as limited substantive gains; from a systemic perspective, it appears as lack of substantive development. Geertz's description emphasizes the individual angle—"people's living standards improved only very slightly"—while Goldenweiser's phenomenon emphasizes the systemic angle: the "definitive form of cultural patterns" could not "transform into new patterns." Huang's (2002) description incorporates both perspectives. Individually, he acknowledged that people could increase income by working longer hours, but due to diminishing marginal product, such gains were ultimately limited. Systemically, he emphasized that this income increase should be distinguished from "development" because the entire system showed no improvement in labor productivity. Synthesizing these analyses, involution can be defined as: a phenomenon of continuously escalating input within an established framework that lacks substantive gains or development.

Does involution exist in education? The answer is likely yes. Limited access to prestigious graduate programs and desirable corporate positions creates a closed space, and competition for these spots constitutes a zero-sum game within this confined space (Chen and Bao, 2022). To avoid losing this zero-sum competition, students must invest more time in studying. However, these investments do not necessarily yield substantive personal benefits. Research based on the China Education Panel Survey found that as extracurricular tutoring time increased overall among middle school students, the effect of individual tutoring time on academic improvement gradually weakened while anxiety and depression increased (Zhang and Gao, 2022). From a learning and growth perspective, excessive study time reduces learning efficiency (Pang and Li, 2010), and academic anxiety easily leads to learning burnout (Gao, 2014). Additionally, Zhang and Gao found that increased tutoring time reduced sleep quality, which harms

brain health and impedes student development (Eide et al., 2021). Moreover, because schools struggle to comprehensively and accurately measure ability and can only summarize performance through grades, some students pursue "scores" rather than "growth." To improve GPA, students may choose courses with low relevance to their major but relatively easy grading (Zhao and Hu, 2022). On many campuses, "GPA-boosting" course selection guides circulate among students. In such cases, even massive time investment does not guarantee substantive knowledge acquisition. From a talent cultivation perspective, the current evaluation system that emphasizes only "credentials" and "prestigious schools" lacks substantive development and undermines the nation's ability to maintain a rational talent cultivation structure (Xiong, 2021). Thus, complaints about educational involution are not unfounded.

Academic involution may have numerous harmful effects. One concern is that university students may become homogenized, losing individuality and creativity (Xia, 2012). Many college students continue an exam-oriented mentality, pursuing scores and advancement rates (Chen, 2019), which involution only exacerbates. The cultivation of high-level talent may also suffer. Under "involuted" doctoral dissertation publication systems, academic papers may exhibit "quantitative growth without substantive innovation" (Zhao, 2021), clearly contradicting the national graduate education conference's requirement to "cultivate high-level innovative talent." Involution may also burden students' mental health and families' educational investments. Severe involution can cause young people to lose learning motivation and may lead to anxiety, depression, and even suicide and violence (Dai, 2022). It may also affect entire families, creating a dilemma of "increasing parental educational investment without reducing anxiety" (Chen and Bao, 2022). More broadly, educational involution wastes social resources and hinders national long-term development.

Although in-depth research on academic involution is imperative, reliable measurement tools remain unavailable. Three studies have attempted to compile academic involution questionnaires (Yan et al., 2022; Yang et al., 2023; Yi et al., 2022), but all suffer from unclear definitions, inadequate content validity, or insufficient reliability and validity evidence. To provide a reliable measurement tool, this study begins by refining the definition of involution and systematically examines the questionnaire's reliability and validity. The operational definition of academic involution is: the phenomenon of escalating input in exam-oriented or ranking competition that lacks growth in knowledge, ability, or personality.

Based on this operational definition, individual academic involution may have two interrelated manifestations: external and internal. Externally, students high in involution engage in excessive input or competition for high GPA or ranking. Internally, they lack growth in ability or personality. Study 1 will construct a representative initial questionnaire based on these manifestations, the definition of involution, literature review, real-world observations, and student interviews. Study 2 will use a large sample to develop the final involution questionnaire and test its validity using zero-sum belief, trait anxiety, life satis-

faction, creative potential, and sleep quality as criterion measures (Wang, 2013). We expect academic involution to correlate positively with zero-sum belief and trait anxiety, and negatively with life satisfaction and sleep quality. Since involution implies lack of substantive development and gains, students high in involution should not show higher creative potential. Study 3 will examine the test-retest reliability of the final questionnaire in a new sample.

---

2. Study 1: Development of the Initial College Students' Academic Involution Questionnaire

2.1 Purpose

Study 1 constructed an initial item pool for the College Students' Academic Involution Questionnaire through integration of the involution definition, real-world observations, and interviews. Quantitative analysis of the initial questionnaire eliminated problematic items, yielding an initial version of the questionnaire.

2.2 Method

2.2.1 Item Development and Questionnaire Structure Based on the operational definition, academic involution likely manifests in two interrelated ways: external and internal. External manifestations involve excessive input or competition aimed at obtaining high GPA or ranking. Internal manifestations involve lack of ability or personality growth. Study 1 constructed the initial item pool by screening real-world examples that reflected these manifestations, drawing upon the literature-based definition of involution, daily observations, and interviews. For example, items reflecting "excessive input" included: "As long as it can improve my GPA ranking, I am willing to invest substantial time in a particular course." Items addressing "excessive competition" included: "I worry about being surpassed by peers while resting." Items reflecting "lack of ability growth" included: "I participate in many extracurricular practices or internships with little real gain, mainly to enrich my resume." Items focusing on "lack of personality growth" included: "I don't have work I truly enjoy; I just try not to fall behind in various competitions." Additionally, a few reverse-scored items were included in the initial pool. For example, a reverse-scored item for "excessive input": "I maintain regular sleep schedules and do not stay up late studying for exams or rankings." For "excessive competition": "Even without competing with others, I would maintain good study habits." For "lack of ability growth": "I strive to make my papers valuable, not just to obtain high scores." For "lack of personality growth": "I choose courses based on my interests, not on how high the instructor grades."

The initial pool contained 31 items, including 8 reverse-scored items. All items used a 5-point Likert scale (1: strongly disagree to 5: strongly agree), with higher scores indicating greater agreement with the item description.

2.2.2 Participants Study 1 administered the questionnaire to current college students. A total of 373 questionnaires were distributed online. Thirty-five participants failed attention checks embedded in the questionnaire and were excluded from analysis. The final sample comprised 338 valid questionnaires (90.62% retention rate). Participants had a mean age of 20.52 ± 1.53 years, including 117 males (34.62%), and represented various academic years: freshman (4.44%), sophomore (29.59%), junior (39.05%), senior (15.98%), and other (10.95%).

2.3 Results

2.3.1 Item Analysis To examine item discrimination, participants were divided into high and low groups based on the 27th percentile of total scores (after reverse-scoring). Independent samples t-tests revealed significant differences between high-scoring (N = 91) and low-scoring (N = 88) groups on 28 items ($p < 0.01$, Bonferroni-corrected). Three items showed non-significant differences (items 7, 12, and 22), indicating poor discrimination.

To examine item heterogeneity, corrected item-total correlations were computed. Based on previous literature, correlations between 0.30 and 0.70 are acceptable (Vagos et al., 2016). Six items fell below 0.30 (items 1, 2, 7, 12, 17, and 22, with correlations of $r_1 = 0.24$, $r_2 = 0.28$, $r_7 = -0.15$, $r_{12} = 0.09$, $r_{17} = 0.23$, $r_{22} = -0.08$), indicating high heterogeneity.

To prevent severely problematic items from affecting subsequent exploratory factor analysis, items with both poor discrimination and high heterogeneity were removed before further analysis (items 7, 12, and 22, all reverse-scored).

2.3.2 Exploratory Factor Analysis Exploratory factor analysis was conducted on the remaining 28 items to preliminarily examine the factor structure. Feasibility statistics showed a KMO value of 0.88, exceeding the empirical standard of 0.80, indicating sufficient common variance among variables. Bartlett's test of sphericity was significant ($^2 = 3174.84$, $df = 378$, $p < 0.001$), confirming suitability for factor analysis.

Principal component analysis extracted factors with eigenvalues greater than 1, followed by Promax oblique rotation. This criterion suggested six factors, explaining 57.14% of total variance. Following previous literature (Chen et al., 2006; Olatunji et al., 2007) and Study 1's purpose (removing severely problematic items while retaining as many acceptable items as possible), factor retention criteria were: (1) item loadings > 0.3 on the factor, and (2) at least three items per factor. All items met criterion (1), but Factor 6 contained only one item (item 14), violating criterion (2) and was considered for deletion. Additionally, five reverse-scored items designed for different dimensions (items 1, 17, 18, 19, and 21) formed a separate factor, failing to serve their intended purpose and were considered for deletion.

2.4 Initial Version of the College Students' Academic Involution Questionnaire

Based on item analysis and exploratory factor analysis, items 1, 7, 12, 14, 17, 18, 19, 21, and 22 were initially slated for removal. However, drawing on previous literature, real-world observations, and interview results, item 1 ("I choose courses based on my interests rather than instructor grading") was frequently mentioned in discussions of involution. Therefore, this item was retained in the initial pool and modified to a forward-scored item: "When selecting courses, I prioritize those where I can easily obtain high grades to improve my GPA," for further examination. Other retained items were slightly modified while maintaining their original meaning to make expressions more concise and direct. Study 1 thus produced an initial version of the College Students' Academic Involution Questionnaire comprising 23 items.

---

3. Study 2: Refinement and Validation of the Questionnaire

3.1 Purpose

Study 2 used a large sample to refine the initial questionnaire and examined the reliability and validity of the revised version in a new sample to establish the final College Students' Academic Involution Questionnaire.

3.2 Method

3.2.1 Participants Study 2 collected data from two samples. Sample 1 was used for exploratory factor analysis to revise the initial questionnaire, while Sample 2 was used for confirmatory factor analysis to test the stability of the revised factor structure. Together, these steps determined the final version of the questionnaire. Based on the final version and criterion data collected from Sample 1, Study 2 further explored relationships between academic involution and zero-sum belief, trait anxiety, life satisfaction, creative potential, and sleep quality to establish structural validity.

Sample 1: The Credamo platform (https://www.credamo.com/#/) was commissioned to collect 3,000 valid questionnaires; all participants passed embedded attention checks. The sample covered 32 provinces, municipalities, and autonomous regions. Participants had a mean age of 20.80 $\pm$ 1.58 years, including 919 males (30.63%), and represented various institution types: Project 985 (11.33%), Project 211 (19.07%), first-tier non-985/211 (29.10%), second-tier (24.90%), third-tier (4.73%), and vocational (9.90%). Academic years included freshman (12.33%), sophomore (23.30%), junior (29.00%), senior (24.70%), and other (10.67%).

Sample 2: University students were recruited independently, with 580 ques-

tionnaires distributed. Nine participants failed attention checks, yielding 571 valid questionnaires (98.44% retention rate). The sample covered 30 provinces, municipalities, and autonomous regions. Participants had a mean age of 21.19 ± 1.73 years, including 166 males (29.07%), and represented institution types: Project 985 (11.91%), Project 211 (15.41%), first-tier non-985/211 (33.10%), second-tier (27.50%), third-tier (5.25%), and vocational (6.30%). Academic years included freshman (10.86%), sophomore (23.47%), junior (31.87%), senior (22.94%), and other (10.86%).

3.2.2 Measures

1. **College Students' Academic Involution Questionnaire:** Sample 1 completed the initial 23-item version. Sample 2 completed the revised 17-item version derived from exploratory factor analysis. Both versions used a 5-point Likert scale (1 = strongly disagree to 5 = strongly agree), with higher scores indicating greater academic involution.

2. **Zero-Sum Belief Scale:** A 6-item scale by Crocker and Canevello (2008) measured zero-sum beliefs (e.g., "One person's success depends on another's failure") using a 7-point Likert scale. Cronbach's α was 0.72 in Sample 1.

3. **Trait Anxiety Scale:** This 20-item scale (e.g., "I feel nervous and uneasy") used a 1-4 rating (1 = not at all to 4 = very much so), with higher scores indicating greater trait anxiety (Wang et al., 1999). Cronbach's α was 0.88 in Sample 1.

4. **Life Satisfaction Scale:** This 5-item scale (e.g., "I am satisfied with my life") used a 7-point Likert scale, with higher scores indicating greater overall life satisfaction (Diener et al., 1985). Cronbach's α was 0.87 in Sample 1.

5. **Williams Creative Tendency Scale:** This 50-item scale assesses four dimensions: curiosity, risk-taking, imagination, and challenge (Liu et al., 2018), including 8 reverse-scored items. After reverse-scoring, higher total scores indicate greater creative potential. Cronbach's α was 0.87 in Sample 1.

6. **Sleep:** Two self-developed items measured sleep duration ("How many hours do you sleep per day on average?" and "Do you feel your sleep is adequate?") and one item assessed insomnia ("Do you have insomnia problems?").

7. **Family Socioeconomic Status:** Subjective family socioeconomic status was measured using the 10-rung MacArthur Scale (Goodman et al., 2001). This information was collected as background data but not included in subsequent analyses.

3.2.3 Common Method Bias Check Exploratory factor analysis was used to test for common method bias. All items from the initial involution questionnaire, zero-sum belief scale, trait anxiety scale, life satisfaction scale, creative tendency scale, and sleep survey completed by Sample 1 were analyzed. The analysis extracted more than one factor, with the first common factor explaining 12.11% of variance, below the 40% threshold, indicating no severe common method bias in Sample 1.

3.3 Results

3.3.1 Item Analysis To examine item discrimination in the initial questionnaire, Sample 1 participants were divided into high and low groups based on the 27th percentile of total scores. Independent samples t-tests revealed significant differences between high-scoring (N = 803) and low-scoring (N = 842) groups on all items ($p < 0.01$, Bonferroni-corrected), indicating good discrimination for all items.

To examine item heterogeneity, corrected item-total correlations were computed. Using the same standard as Study 1, three items showed correlations below 0.30 (items 2, 13, and 29, with $r_2 = 0.17$, $r_{13} = 0.28$, and $r_{29} = 0.19$), indicating high heterogeneity.

To ensure high reliability and validity, items failing to meet acceptable standards in Study 2 were considered poor reflections of academic involution and slated for deletion. Therefore, before subsequent analysis, three heterogeneous items (2, 13, and 29) were removed.

3.3.2 Exploratory Factor Analysis KMO and Bartlett's tests were conducted on the remaining 20 items. Results showed a KMO value of 0.91, exceeding the 0.80 standard, and Bartlett's test was significant ($^2 = 14507.78$, df = 190, $p < 0.001$), confirming suitability for factor analysis.

Principal component analysis extracted factors with eigenvalues greater than 1, with the number of factors determined by both eigenvalue and scree plot criteria, followed by Promax rotation. Both criteria suggested three factors (factor loadings shown in Table 1). The three-factor model explained 43.69% of total variance. Following previous literature (Chen et al., 2006) and Study 2's purpose (ensuring reliability and validity while retaining well-performing items), factor retention criteria were: (1) item loadings > 0.4, and (2) at least three items per factor. All factors met criterion (2), but three items (3, 5, and 26) had loadings below 0.4 on any factor, violating criterion (1) and were considered for deletion.

The retained 17 items distributed across three factors. Factor 1 contained 7 items and was named "Unwilling Hardworking"—pursuing external outcomes while losing intrinsic learning motivation. Factor 2 contained 5 items and was named "Excessive Competition"—over-competing due to limited resources. Factor 3 contained 5 items and was named "Surface Learning"—pursuing high GPA

rather than genuine scholarship. This 17-item revised questionnaire was tested for factor structure stability in Sample 2.

Table 1 Factor Loading Matrix of the College Students' Academic Involution Questionnaire

Item	Unwilling Hardworking	Excessive Competition	Surface Learning
4. Apart from exam preparation, I don't know what use my knowledge is.			
6. I prefer doing practice problems over reading or thinking about theory because it improves scores better.			

Item	Unwilling Hardworking	Excessive Competition	Surface Learning
15. Although I appear to be working hard, I don't seem to learn much.			
23. I study hard, but not because I truly love learning.			
25. I don't have work I truly enjoy; I just try not to fall behind in competitions.			

Item	Unwilling Hardworking	Excessive Competition	Surface Learning
27. After crossing the threshold of 'outperforming competitors,' I stop any additional learning.			
31. We work harder and harder, but life doesn't seem to get better.			

Item	Unwilling Hardworking	Excessive Competition	Surface Learning
16. I worry about being surpassed by peers while resting.			
20. I'm busy improving competitiveness and have no time for what I truly want to do.			

Item	Unwilling Hardworking	Excessive Competition	Surface Learning
24. If I find classmates studying while I'm resting, I feel very guilty.			
28. Temporary leads or lags in competition frequently affect my mood and study state.			
30. I study exhaustingly but rarely have time to rest.			

Item	Unwilling Hardworking	Excessive Competition	Surface Learning
1. When selecting courses, I prioritize those where I can easily obtain high grades to improve my GPA.			

Item	Unwilling Hardworking	Excessive Competition	Surface Learning
8. For extra credit (course, graduate recommendation, etc.), I participate in activities with little substantive meaning that are merely formalities.			

Item	Unwilling Hardworking	Excessive Competition	Surface Learning
9. I participate in many internships or practices with little real gain, mainly to enrich my resume.			
10. When completing course papers, I intentionally exceed word limits to obtain better grades.			

Item	Unwilling Hardworking	Excessive Competition	Surface Learning
11. As long as it can improve my GPA ranking, I'm willing to invest substantial time in a particular course.			

Note: Factor loadings below 0.4 are not shown. Item 24 was deleted during confirmatory factor analysis in Sample 2.

3.3.3 Confirmatory Factor Analysis To test the stability of the revised questionnaire's factor structure from Sample 1, confirmatory factor analysis was conducted using Sample 2. Initial results showed strong residual correlations between items 16 and 24 (modification index = 41.97). Content analysis revealed these items overlapped considerably, both measuring whether excessive competition affects rest (Item 16: "I worry about being surpassed by peers while resting"; Item 24: "If I find classmates studying while I'm resting, I feel very guilty"). The correlation between items 16 and 24 was strong ($r = 0.54$, $p < 0.001$, 95% CI = [0.48, 0.60]). Therefore, the item with lower factor loading (item 24) was deleted.

Confirmatory factor analysis on the remaining 16 items yielded acceptable model fit: $^{2} = 353.45$, df = 101, GFI = 0.93, RMSEA = 0.07, CFI = 0.90. These indices indicated that the 16-item, three-factor structure was stable, establishing the final College Students' Academic Involution Questionnaire.

3.3.4 Internal Consistency Analysis Internal consistency of the final questionnaire was analyzed using Sample 2 data. All three subscales correlated significantly with the total score (r_{{{unwilling}}*{{hardworking}}} *= 0.85, p < 0.001, 95% CI = [0.82, 0.87]; r*{{{excessive}}*{{competition}}} *= 0.80, p < 0.001, 95% CI = [0.77, 0.83]; r*{{{surface}}*{{learning}}} = 0.73, p < 0.001, 95% CI = [0.69, 0.77]). Subscales intercorrelated strongly (r*{{{unwilling}}*{{hardworking}}}*-excessive*{competition} = 0.49, p < 0.001, 95% CI = [0.43, 0.55]; r_{{{unwilling}}*{{hardworking}}}*-surface*{learning} = 0.34, p < 0.001, 95% CI = [0.27, 0.41]; r_{{{excessive}}*{{competition}}}*-surface*{learning} = 0.54, p < 0.001, 95% CI = [0.48, 0.60]). Cronbach's α for the total questionnaire was 0.86, and for the three subscales were 0.83, 0.72, and 0.70, respectively—all acceptable. These results demonstrate good internal consistency and item reliability for the final questionnaire.

3.3.5 Structural Validity Based on the final questionnaire structure, total involution scores and three subscale scores were computed for Sample 1 and correlated with zero-sum belief, anxiety, life satisfaction, creative potential, and sleep. Results (see Table 2) showed that relationships between total involution scores and criterion variables matched expectations: high involution was associated with stronger zero-sum beliefs, higher anxiety, lower life satisfaction, shorter sleep duration, less adequate sleep, and more severe insomnia, but not higher creative potential.

Bayesian statistics were used to test whether the null hypothesis of no relationship between academic involution and creativity could be accepted. The Bayes factor $BF_{01} = 10.40$ indicated that the data were 10.40 times more likely under the null hypothesis than the alternative. According to classification criteria (Jeffreys, 1961), this provides moderate evidence supporting the null hypothesis. Relationships between the three subscales and most criterion variables showed patterns similar to the total score, but their relationships with creative potential diverged markedly. Individuals high in unwilling hardworking showed lower creative potential, whereas those high in excessive competition and surface learning showed higher creative potential. Additionally, compared to unwilling hardworking and excessive competition, surface learning showed weaker correlations with anxiety, sleep duration, and insomnia, and near-zero correlations with sleep adequacy and life satisfaction (sleep adequacy: $BF_{01} = 12.25$, strong evidence for the null; life satisfaction: $BF_{01} = 34.79$, strong evidence for the null). Overall, relationships between the academic involution questionnaire and its subscales with criterion variables align with the behavioral pattern of high involution individuals who over-compete, focus on grades or rankings, and neglect substantive personal growth, demonstrating good structural validity for the final questionnaire.

Table 2 Structural Validity of the Final College Students' Academic Involution Questionnaire

Variable	M (SD)	Total	Unwilling Hardworking	Excessive Competition	Surface Learning
1. Zero-sum belief	23.21 (6.46)	0.38*** [0.35, 0.41]	0.35*** [0.32, 0.38]	0.31*** [0.28, 0.34]	0.23*** [0.19, 0.26]
2. Trait anxiety	55.47 (9.27)	0.40*** [0.37, 0.43]	0.45*** [0.42, 0.47]	0.38*** [0.35, 0.41]	0.07*** [0.04, 0.11]
3. Life satisfaction	19.16 (6.63)	-0.23*** [-0.27, -0.20]	-0.31*** [-0.34, -0.28]	-0.17*** [-0.21, -0.14]	-0.01 [-0.05, 0.02]
4. Creative potential	173.70 (18.56)	-0.03 [-0.07, 0.00]	-0.20*** [-0.23, -0.16]	0.07*** [0.04, 0.11]	0.14*** [0.11, 0.18]
5. Sleep: Duration	-	-0.11*** [-0.15, -0.08]	-0.06*** [-0.09, -0.02]	-0.17*** [-0.21, -0.14]	-0.05** [-0.08, -0.01]
6. Sleep: Adequacy	3.37 (0.83)	-0.16*** [-0.20, -0.13]	-0.15*** [-0.18, -0.11]	-0.21*** [-0.25, -0.18]	-0.03 [-0.06, 0.01]
7. Sleep: Insomnia	-	0.12*** [0.08, 0.15]	0.10*** [0.06, 0.13]	0.14*** [0.11, 0.18]	0.05* [0.01, 0.08]

Note: $p < 0.05$. ** $p < 0.01$. *** $p < 0.001$. Values in brackets represent 95% confidence intervals for correlation coefficients. Sleep duration and insomnia are categorical variables using Spearman correlation.*

3.4 Final Version of the College Students' Academic Involution Questionnaire

Study 2 established the final College Students' Academic Involution Questionnaire based on item analysis, exploratory factor analysis, and confirmatory factor analysis. The final questionnaire contains 16 items across three subscales: unwilling hardworking, excessive competition, and surface learning. Further internal consistency and structural validity analyses demonstrated good reliability and validity for both the total questionnaire and its subscales.

4. Study 3: Test-Retest Reliability

4.1 Purpose

Study 3 administered the final questionnaire to a new sample with a one-month interval to examine test-retest reliability.

4.2 Method

4.2.1 Participants One hundred twelve undergraduate students participated in the retest. Six failed to complete the questionnaire within the one-month interval, and seven failed attention checks. The final sample comprised 99 valid participants (29 males, 29.29%), including freshmen (57), sophomores (11), juniors (30), and seniors (1).

4.2.2 Measure Final version of the College Students' Academic Involution Questionnaire.

4.3 Results

The one-month test-retest reliability for the final questionnaire was 0.83 ($p < 0.001$, 95% CI = [0.76, 0.89]). Subscale retest reliabilities were 0.79 ($p < 0.001$, 95% CI = [0.71, 0.86]), 0.77 ($p < 0.001$, 95% CI = [0.68, 0.84]), and 0.80 ($p < 0.001$, 95% CI = [0.71, 0.86]), all exceeding 0.70, indicating good test-retest reliability.

Internal consistency was analyzed for both test and retest data. In the test data, Cronbach's α was 0.84 for the total questionnaire and 0.77, 0.71, and 0.71 for the three subscales, all exceeding 0.70. In the retest data, Cronbach's α values were 0.89 for the total questionnaire and 0.85, 0.74, and 0.78 for the subscales, again exceeding 0.70. These results demonstrate good item reliability for the final questionnaire.

5. General Discussion

This study developed a 16-item, three-factor College Students' Academic Involution Questionnaire based on a literature-derived definition of involution, systematic item development, and comprehensive reliability and validity testing. Large-sample results demonstrated reasonable criterion-related validity: students with higher academic involution scores showed stronger zero-sum beliefs, higher trait anxiety, lower life satisfaction, poorer sleep quality, but not higher creative potential. The three factors were named unwilling hardworking (pursuing external outcomes while losing intrinsic learning motivation), excessive competition (over-competing due to limited resources), and surface learning

(pursuing high GPA rather than genuine scholarship). Notably, Yuan and Xing's (2021) description of college student involution aligns remarkably well with our three-factor structure, providing content-level support for the questionnaire's validity.

5.1 Theoretical Implications

This study deepens understanding of the involution concept. Based on Goldenweiser, Geertz, and Huang's descriptions, we identified three core characteristics: (1) constraints of an established framework, (2) continuously escalating input, and (3) lack of substantive gains or development. These characteristics are compatible with various descriptions of involution. For example, Liu and Qiu (2004) "restricted external expansion" and Ji (2010) "external constraints, self-locking" reflect established framework constraints. Fan (2004) "increasing internal refinement" and Wang (2013) "self-repetition" represent forms of escalating input within established frameworks. Xia (2012) "growth without development" and Zhao (2021) "quantitative growth without substantive innovation" reflect lack of substantive gains and development. Chen and Bao (2022) summary of involution characteristics also yielded three similar features: (1) limited space, (2) increased input, and (3) failure to achieve win-win outcomes individually or collectively. Logically, all three core characteristics are essential. "Escalating input without substantive gains" represents people's intuitive experience of involution, while "established framework constraints" constitute the prerequisite condition. Thus, defining involution as "a phenomenon of continuously escalating input within an established framework that lacks substantive gains or development" offers both theoretical soundness and practical utility.

5.2 Practical Applications

This study provides a more reliable measurement tool for future involution research. Previous academic involution questionnaires suffer from unclear definitions, inadequate content validity, or insufficient reliability and validity evidence (Yan et al., 2022; Yang et al., 2023; Yi et al., 2022). Yi et al. defined involution as "irrational competition for limited resources" but did not require that such competition fails to produce substantive gains. For example, their item "I study in my free time because I want to acquire more knowledge to improve my competitiveness" suggests individuals can gain substantive knowledge while the system improves through healthy competition. Yan et al. emphasized "increasing burden and excessive competition in irrational environments," with items like "I anxiously study to keep up with others' pace" and "I feel uneasy when I don't do what most people do," which do not align with the involution concept. Yang et al. neither clearly defined involution nor provided reliability and validity evidence, with many items bearing little relation to the concept (e.g., "I follow knowledge-sharing accounts on social media to improve myself so I don't fall behind" and "I wake up early and return late to study so I don't fall behind"). While Yi and Yang both view involution as "irrational competition,"

they fail to define "irrational." Xu (2021) criticized involution's overly broad reference, arguing it should not "refer to any excessive competition" and calling for strict criteria to delimit the concept's boundaries and preserve its utility. In response, our questionnaire, based on a clearly defined concept with ample reliability and validity evidence, provides a more useful tool for standardizing involution research.

5.3 Real-World Significance

This study demonstrates that college student academic involution warrants serious attention from educators and society. Nationwide data from 3,000 participants show that academic involution not only exists but is quite prevalent. Using a cutoff of total score > 48 (mean item rating > 3, indicating "agree" or "strongly agree" for most items), 1,930 individuals (64.33%) qualified as involuted. Using the same criterion, 61.03% exhibited unwilling hardworking, 42.83% showed excessive competition, and 71.33% demonstrated surface learning. Negative effects were also evident. High involution individuals had higher anxiety, lower life satisfaction, and poorer sleep quality. Higher unwilling hardworking was also associated with lower creative potential, contrary to higher education's goals. Although excessive competition might promote creative potential, this likely comes at the cost of sleep. Since sleep deprivation impairs memory and immunity, this effect is unsustainable and may increase individual and societal healthcare costs (Gais et al., 2006; Prather et al., 2015). Students high in surface learning also showed higher creative potential but may struggle to achieve creative accomplishments, which require not only creative tendency but also tenacious determination and endless inquiry (Zhang et al., 2020)—qualities surface learners likely lack. In academics, surface learners' creativity may manifest as finding "shortcuts" to better grades. Indeed, data from 3,000 participants showed that high surface learning was associated with better class ranking ($| | = 0.14$, $p < 0.001$), corroborating Yuan and Xing's (2021) "evaluation distortion" claim and warning educators that current evaluation systems may inadvertently "cultivate" surface learners and raising questions about their societal impact after graduation.

5.4 Limitations and Future Directions

While this study advances involution research, three important issues remain unresolved. First, should involution be judged by objective or subjective criteria? Sociologists prefer objective marginal efficiency criteria, but this is often impractical, especially for aspects dependent on subjective experience (e.g., learning gains). For applicability, our questionnaire uses subjective experience, assuming it sensitively reflects objective systemic conditions. Second, the boundary between involution and non-involution is unclear. Huang (2002) proposed that involution occurs when marginal product begins to decrease, but Liu and Qiu refuted this as inconsistent with Geertz's definition and unrealistic. This difficulty is especially pronounced in academic involution; no widely accepted di-

viding line exists. Our questionnaire only distinguishes high and low involution individuals, leaving cutoff determination for future research. Third, involution's opposite remains undefined. Our definition incorporates three widely accepted characteristics, but should involution's opposite oppose all three characteristics or only some? Reverse-scored items in our initial version attempted to measure "no clear closed space, escalating input but with substantive gains," but failed to function as expected. What should constitute involution's opposite? Is "lying flat" a more suitable opposite? These questions await future exploration.

6. Conclusion

This study developed a 16-item College Students' Academic Involution Scale comprising three dimensions: unwilling hardworking, excessive competition, and surface learning. In a nationwide sample of 3,000 participants, over 60% of college students identified with involution descriptions. Students scoring higher on the scale showed stronger zero-sum beliefs, higher trait anxiety, lower life satisfaction, poorer sleep quality, but not higher creative potential. These findings demonstrate that academic involution is both prevalent and demonstrably harmful in college education, meriting serious attention from society and educators.

References

Chen, C., & Bao, L. (2022). The Origin of Involution and Solutions to Address Involution in Education. Journal of China Examinations, 2, 81-88.

Chen, H., Jackson, T., & Huang, X. (2006). The Negative Physical Self Scale: Initial development and validation in samples of Chinese adolescents and young adults. Body Image, 3(4), 401–412.

Crocker, J., & Canevello, A. (2008). Creating and undermining social support in communal relationships: The role of compassionate and self-image goals. Journal of Personality and Social Psychology, 95(3), 555–575.

Diener, E., Emmons, R. A., Larsen, R. J., & Griffin, S. (1985). The Satisfaction with Life Scale. Journal of Personality Assessment, 49(1), 71–75.

Eide, P. K., Vinje, V., Pripp, A. H., Mardal, K.-A., & Ringstad, G. (2021). Sleep deprivation impairs molecular clearance from the human brain. Brain, 144(3), 863–874.

Gais, S., Lucas, B., & Born, J. (2006). Sleep after learning aids memory recall. Learning & Memory, 13(3), 259–262.

Gao, M. (2014). Mediating Effect of Academic Emotion between College Adjustment and Learning Burnout. Chinese Journal of Clinical Psychology, 22(3), 537–539.

Geertz, C. (1963). Agricultural involution: The processes of ecological change in Indonesia. University of California Press.

Goodman, E., Adler, N. E., Kawachi, I., Frazier, A. L., Huang, B., & Colditz, G. A. (2001). Adolescents' perceptions of social status: Development and evaluation of a new indicator. Pediatrics, 108(2), 1–8.

Guo, J. Q. (2007). New Understanding of the "Involution" Concept. Sociological Studies, 3, 194–208.

Huang, Z. Z. (2021). Re-thinking Involution and De-involution. Open Times, 1, 157–168.

Huang, Z. Z. (2002). Development or Involution? 18th Century Britain and China—A Review of Pomeranz's "The Great Divergence: Europe, China, and the Making of the Modern World Economy". Historical Research, 4, 150–162.

Jeffreys, H. (1961). The theory of probability (3rd ed). Oxford University Press.

Ji, Y. P. (2010). Theoretical Review of "Involution". Journal of Changchun University of Technology (Social Sciences Edition), 22(3), 48–49.

Liu, S. D., & Qiu, Z. Q. (2004). Conceptual Analysis of "Involution". Sociological Studies, 5, 96–110.

Liu, Y., Wu, B., Petti, C., Wu, X., & Han, S. (2018). Self-construals moderate associations between trait creativity and social brain network. Neuropsychologia, 111, 284-291.

McCullough, C. (2019). Review of "agricultural involution: The processes of ecological change in Indonesia" by Clifford Geertz. International Journal of Anthropology and Ethnology, 3, 1–5.

Olatunji, B. O., Williams, N. L., Tolin, D. F., Abramowitz, J. S., Sawchuk, C. N., Lohr, J. M., & Elwood, L. S. (2007). The Disgust Scale: Item analysis, factor structure, and suggestions for refinement. Psychological Assessment, 19(3), 281–297.

Pang, S. X., & Li, S. X. (2010). Involution: A Description of Bottleneck in Education. Higher Education Development and Evaluation, 26(6), 24–29.

Prather, A. A., Janicki-Deverts, D., Hall, M. H., & Cohen, S. (2015). Behaviorally assessed sleep and susceptibility to the common cold. Sleep, 38(9), 1353–1359.

Vagos, P., Rijo, D., & Santos, I. M. (2016). Scenes for Social Information Processing in Adolescence: Item and factor analytic procedures for psychometric appraisal. Psychological Assessment, 28(4), 416–428.

Wang, C. (2013). Involution and its Cultural Psychological Mechanism. Journal of Shenzhen University (Humanities & Social Sciences), 5, 182–190.

Wang, X. D., Wang, X. L., & Ma, H. (1999). Mental Health Assessment Scale Manual (Revised Edition, pp. 235-238). Beijing: Chinese Mental Health Journal.

Xia, J. (2012). Discussion on Involution Growth in Higher Education. Education Review, 3, 12–14.

Xiong, B. Q. (2021). Could Implementing Ten-Year Compulsory Education Alleviate "Involution"? Online Learning, 8, 79.

Xu, Y. J. (2021). Digital Fetishism: The Essence of Involution. Exploration and Free Views, 3, 57–65.

Yan, D., Zhang, H., Guo, S., & Zeng, W. (2022). Influence of anxiety on university students' academic involution behavior during COVID-19 pandemic: Mediating effect of cognitive closure needs. Frontiers in Psychology, 13.

Yang, Y., Peng, Y., Li, W., Lu, S., Wang, C., Chen, S., & Zhong, J. (2023). Psychometric evaluation of the academic involution scale for college students in China: An application of Rasch analysis. Frontiers in Psychology, 14.

Yi, D., Wu, J., Zhang, M., Zeng, Q., Wang, J., Liang, J., & Cai, Y. (2022). Does Involution Cause Anxiety? An Empirical Study from Chinese Universities. International Journal of Environmental Research and Public Health, 19(16), 9826.

Zhang, Q., & Gao, Y. Y. (2022). Competition and Game: The Academic Returns and Mental Health Costs of Private Tutoring. Chinese Journal of Sociology, 42(3), 159–194.

Zhang, Z. S., Hoxha, L., Aljughaiman, A., Arnliu, A., & Ziegler, A. (2020). Social Environmental Factors and Personal Motivational Factors Associated with Creative Achievement: A Cross-cultural Perspective. The Journal of Creative Behavior, 54(1), 211-225.

Zhao, T. T., & Hu, Y. T. (2022). The Ability and Response: The Function of the Grade Point Average System from the Undergraduate Transcripts. University Education Science, 3, 22–30.

Zhao, X. H. (2021). "Involution" of the Doctoral Publication System: The Characterization, Mechanism and Correction. Journal of Higher Education Management, 15(3), 104–113.

Chen, X. F. (2019). Analysis of College Students' Learning Psychology under Exam-Oriented Education Background. Campus Psychology, 17(3), 210–212.

Dai, Z. H. (2022). Cultivating Playful Spirit: A Way to Break Through Educational Involution. Youth and Children Studies, 1, 66–72.

Fan, Z. H. (2004). On the "Involution" Problem in China's Institutional Innovation. Society, 4, 4–7.

ChinaRxiv 中X

Yuan, J. S., & Xing, T. Q. (2021). “Entering and Breaking the Game”: The Behavioral Logic and Self-Rescue of College Student Involution Participants. Higher Education Exploration, 10, 123–128.

Note: Figure translations are in progress. See original paper for figures.

Source: ChinaXiv — Machine translation. Verify with original.