**Title**

Exploring the Ontology of Self-Enhancement: A Data-Driven Perspective

**Abstract**

**Keywords**

**1 Introduction**

1.1 Research Background

Self-enhancement is......

Its significance in psychology......

Although there is much research on self-enhancement, there are still limitations such as lack of clarity in definitions and measurement methods......

1.2 Research Purpose

Dedicated to data-driven discovery of the ontology of self-enhancement and clarification of the definition of self-enhancement .

Explore which indicators of self-enhancement are highly predictive of psychological adjustment and apply the results to mental health .

Refine the measurement of the self-enhancement domain from multiple perspectives including self-report, criterion comparison, and cognitive tasks .

1.3 Research Hypotheses

Self-enhancement consists of distinct dimensions (e.g., narcissism vs. self-promotion).

These dimensions predict psychological outcomes such as depression and subjective well-being. Different prediction effects at different dimensions .

2 Results

2.1 Association between tasks and questionnaire

The Pearson correlation analysis of self-report questionnaires revealed significant positive correlations among most scales, suggesting a consistent measurement of a common latent construct—self-enhancement. However, two scales, the Hypersensitive Narcissism Scale (HSNS) and the Narcissistic Personality Inventory (NPI), exhibited distinctive correlation patterns that warrant further exploration. Specifically, the HSNS demonstrated negative correlations with other scales, which may reflect the tendency of hypersensitive narcissistic individuals to display vulnerability and insecurity. These characteristics might conflict with the positive self-representations assessed by other scales. In contrast, the NPI showed near-zero correlations with other scales, indicating that overt narcissism represents a unique pattern of self-enhancement. Notably, apart from the HSNS and NPI, the other self-report questionnaires exhibited moderate-to-strong positive correlations. This consistency suggests that these scales effectively capture different dimensions of self-enhancement. Taken together, the findings indicate that the selected self-report scales comprehensively reflect participants' levels of self-enhancement. Moreover, the observed correlation patterns provide valuable insights into the multidimensional structure of self-enhancement, highlighting the potential for further theoretical and empirical investigation.

2.2 Dimensionality reduction analysis

2.3 Exploratory factor analysis

2.4 Cluster analysis

2.5 Bifactor analysis

2.6 Prediction analysis

**3 Methods**

3.1 Data source and reliability

The data for this study came from a nationwide online experiment designed to measure individual levels of self-enhancement, obtain individual self-enhancement data, and make data-driven discoveries about the ontology of self-enhancement. One of the errors in online experiments is the uniformity and standardization of the experimental equipment, so the experimenter recorded the participants' web page refresh rate and resolution to reduce random errors due to the experimental design. Adult participants between the ages of 18 and 59 were recruited for the study, with a total of 771 participants participating in the experiment, and a total of 503 participants (mean age = 24 years, SD = 6 years) completed the experimental tasks and questionnaires from Day 1 through Day 4 after screening. These participants met the following screening criteria:

Behavioral data screening:

1. subject data from the 4 blocks of the joint task of the implicit associative reference test were removed for trials with reaction times greater than 10,000 ms and more than 10% of trials with reaction times less than 300 ms.

2. the percentage of correctness in the associative learning paradigm was calculated for the individual graphic conditions and for each domain, and if the percentage of correctness in any graphic condition was less than 20% or in any domain was less than 60%, the subject data were removed.

3. Subjects whose rate of correct recongnition was lower than the random level (55%) were excluded from the old and new vocabulary judgment phase of the self-reference paradigm; also, if more than 10% of the trials in the source judgment phase had a response time of less than 200 ms, that subject was also excluded.

Questionnaire data screening:

Subjects with abnormal responses to the attention detection questions inserted in the questionnaire on days 2, 3, and 10 were excluded.

In terms of grouping the participants, the study divided the participants into young (18-44 years) and middle-aged (45-59 years) groups based on the United Nations World Health Organization's age segmentation criteria. Meanwhile, the monthly household income was grouped based on the 2019 household income report of the National Bureau of Statistics.

3.2 Measures

Participants completed a battery of self-report questionnaires and behavioral tasks (14 indicators from SRET, IAT, and ALT). This study used a series of measurement tools to provide a comprehensive picture of participants' levels of self-enhancement from both implicit and explicit perspectives, as well as the Life Satisfaction Scale, Depressive Disorder Scale (PHQ-9), Generalized Anxiety Scale (GAD-7), and a shortened version of the General Procrastination Scale to measure participants' levels of subjective well-being, depression, anxiety, and procrastination as a way of examining the self-enhancement's ability to real-world outcome Predictive ability.

3.2.1 Explicit measurement

For self-enhancement in the moral domain, the study used the Moral Identity Scale and the Moral Self-Image Scale to measure it. Identity with the concept of moral traits was measured using the Chinese version of the Revised Moral Identity Scale, in which moral identity can be divided into implicit and explicit dimensions, which reflect, respectively, the self's internal identification with moral traits and whether or not the traits are manifested in outward behaviors (Wan, 2008). The scale was derived from the Moral Identity Scale developed by Aquino (2002) et al.

The scale was derived from the moral identity scale developed by Aquino (2002) and others, which has good reliability and validity, with Cronbach's α = 0.83. The revised scale has a good reliability and validity, with Cronbach's α = 0.83.

The revised Chinese version has a Cronbach's α = 0.85, with implicit dimension Cronbach's α = 0.83 and explicit dimension Cronbach's α = 0.74. The Chinese version of the Moral Identity Scale consists of 10 questions, of which 1, 2, 4, 7, and 10 are implicit dimensions, and 3, 5, 6, 8, and 9 are explicit dimensions, with a 5-point scale. Moral self-image was measured using the moral self-image scale translated by Liu, Qinglan et al. (2020) (Jordan et al., 2015), the Chinese version of which has a Cronbach's α = 0.88, with nine nine-point scoring entries, asking participants to judge the extent to which statements about moral self-image correspond to their own. The scale consists of nine 9-point items asking the respondents to judge the extent to which the statements about moral image are consistent with their own, with a total score ranging from 9 to 81. The specific scoring format is as follows: 1 means that it is far from where the respondent wants to be; 5 means that it is exactly the same as where the respondent wants to be; and 9 means that it is far more than where the respondent wants to be. Higher scores indicate higher levels of perceived moral self-image.

In addition, the Rosenberg Self-Esteem Scale developed by Rosenberg (1965) was used in this study. The revised Chinese version of the Self-Esteem Scale (Sun, 2007) was used in this study, which consists of 10 items with 4 points, including 5 reverse scores and 5 positive scores, and is used to assess the overall feelings of self-worth and self-acceptance, and the higher the score, the higher the level of self-esteem, with a Cronbach' s alpha coefficient of 0.5. s α coefficient is 0.835, and the retest reliability is 0.655.

In this paper, the translated Chinese version of the Core Self- Evaluations scale (Judge et al., 2003) was used to measure individual self-evaluations (Du, 2012), which reflects the participants' most basic evaluations of their own abilities and values. The scale has a Cronbach's alpha coefficient of 0.83 and a split-half reliability of 0.84, showing good internal consistency reliability. The scale has a total score range of 10 - 50, with 10 5-point scores, 4 positive and 6 negative scores, with higher scores indicating higher levels of core self-evaluation.

The study referred to the Self concept clarity scale developed by Campbell et al. (1996), and used the translated Chinese version of the scale (Niu, 2016), the internal consistency coefficients of the Chinese version of the scale, Cronbach's coefficient of agreement, and Cronbach's coefficients of agreement.

The internal consistency coefficient of the Chinese version of the scale was Cronbach's α of 0.81. The scale was a 5-point scale with 12 questions, of which 6 and 11 were positively scored, and the rest of the questions were negatively scored. The higher the total score, the higher the clarity of self-concept.

The revised Chinese version of the Life Orientation Test-Revised (LOT-R) (Juanjuan Wen, 2007) was used in the study. The LOT-R consists of 6 items with a 5-point scale divided into optimistic and pessimistic dimensions, and the pessimistic dimensions were reverse scored and added to the optimistic dimensions to obtain the total optimistic personality score. The internal consistency coefficient was 0.78 and the retest reliability was 0.79.

Two subscales of the Balanced Expectancy Response Questionnaire (BIDR) developed by Paulhus (1988) were used in the study, Self-Deception Exaltation (SDE) and Impression of Manipulation (IM). The Self-Deception Exaltation Scale (SDE) has a Cronbach's alpha coefficient of 0.68 - 0.80 (Wang et al., 1999), which reflects the egocentric tendency of participants to make overly positive self-statements. The Impression of Manipulation Scale (IM) had Cronbach's alpha coefficients of 0.75 - 0.86 (Xiangdong Wang et al., 1999) and measured positive moral bias in self-perception.

The internal control subscale (I) of the Internal Control, Authority, and Opportunity Control Orientation Scale (IPC) developed by Levenson (1981) was used in the study, which consisted of 8 items scored on a 7-point scale, with an additional 24 points added to the original total score and a range of 0-48, to measure individuals' belief in their ability to control their own affairs. - 48 measures the extent to which individuals believe they are in control of their lives. The internal locus of control subscale reliability was 0.64 (Wang, 1999) and the retest reliability was 0.8 (Xiao, Li & Chen, 1989).

For the narcissism measure, the study used the Chinese version (Wang, 2008) of the Explicit Narcissism Scale (Ames et al., 2006), which consists of 16 multiple-choice questions, forced-choice, in which the participant is required to choose a more appropriate description of himself/herself from the two choices in each question, covering both self-appraisal and behavioral styles, and the score range is from 0 to 16 points. The scale has an internal consistency reliability of Cronbach's α = 0.71, and the higher the score, the higher the individual's level of dominant narcissism. Meanwhile, the study used the Hypersensitivity Narcissistic Scale (HNS), which is a one-factor structured self-report scale containing 10 items with a 5-point scale regarding self-evaluation and behavioral tendencies of the self, reflecting oversensitivity and vulnerability. The Cronbach's alpha coefficient was 0. 76. Full scale scores represent the implicit dimensions of narcissism, with scores ranging from 10 to 50, with higher scores reflecting higher levels of implicit narcissism.

The Domain Self-Rating Scale (MacDonald et al., 2003) was used to measure subjects' social beliefs about the self in specific domains. The scale consists of five entries in the domains of ability, physical attractiveness, material wealth, social competence, and morality, and subjects were asked to compare themselves with their peers to determine which level they were at. The rating scale is divided into 1 to 12, with 1 representing very low and 12 representing very high, and the reliability of this questionnaire is 0.76 (Hu Chuanpeng, 2017).

In this study, subjective socioeconomic status was measured using the Subjective Socioeconomic Scale (SSS), which has two entries and uses a ten-level scale to measure the family's economic status in the overall social environment and its socioeconomic status in the school environment, respectively. Subjects were asked to determine where their family's economic level stood in society as a whole and where it stood in comparison to their surrounding classmates. For each entry, 1 represents the lowest level of subjective socioeconomic status and 10 represents the highest level of subjective socioeconomic status, with higher scores representing higher subjective socioeconomic status. The re-test reliability of the scale was 0.75, of which the re-test reliability of the subjective economic status of the family was 0.76 and the re-test reliability of the subjective economic status of the school was 0.71 (Hu, Mouli et al., 2012).

3.2.2 Implicit measurement

The implicit tests studied used the common self-referential memory paradigm, the implicit association test, and the associative learning paradigm as implicit measurement tools for self-enhancement.

The vocabulary for the self-referential memory experiment was selected from (Lulu et al., 2023). The self-referential memory experiment used 160 vocabulary words analyzed in the formal experiment with 8 distractor words and 8 practice words. There were 88 trait words for each of the moral and ability domains, and half of the positive and half of the negative words.

The selection of target words for the implicit association test was based on previous research (Cai, 2003), with 6 target words for self and 6 target words for others, and 24 trait words (6 positive morals, 6 negative morals, 6 positive abilities, and 6 negative abilities) from a vocabulary list consisting of the 176 words mentioned above. The associative learning paradigm uses 4 simple geometric shapes: equilateral triangle, square, rhombus, trapezoid, circle, pentagon, ellipse, and hexagon.

3.3 Data Analysis Methods

3.3.1 Correlation Analysis

Correlation analysis is a statistical technique used to examine the strength and direction of the relationship between two variables. In this study, Pearson correlation analysis was primarily used. The Pearson correlation coefficient (r) is the most commonly employed measure for assessing linear relationships between continuous variables, quantifying the degree of linear association between them.Pearson correlation is widely applied in psychology, social sciences, and natural sciences to explore relationships between variables and to generate hypotheses for further research.

In this study, the Pearson correlation coefficient was used to analyze the relationships among self-report questionnaire data, behavioral indicators, and between self-report data and behavioral indicators. The goal was to preliminarily uncover potential connections between various measurement indices of self-enhancement.

3.3.2 Exploratory Factor Analysis (EFA)

EFA was conducted on the combined dataset to explore the underlying structure of the questionnaire and behavioral task data. The number of factors was determined using parallel analysis, and the "minres" extraction method was employed with "oblimin" rotation to allow for correlated factors. Factor loadings exceeding a threshold of 0.3 were considered significant.

In this study, dimensionality reduction analysis was performed on the collected self-report questionnaire and behavioral data. The results of the analysis served as evidence for the potential multidimensional structure of self-enhancement ontology.

3.3.3 Bifactor Analysis

To further evaluate the hierarchical structure of self-enhancement, a bifactor model was employed. The bifactor model is a type of factor analysis model designed to assess the internal structure of a broad construct. Compared to other factor analysis models, the bifactor model is characterized by its ability to simultaneously estimate a general factor and multiple specific factors. General Factor: Represents the common characteristics or overall trend of the construct being measured. For instance, in the context of self-enhancement, the general factor reflects the overarching tendency toward "overall self-enhancement." Specific Factors: Represent unique, specific variances within each subdomain or aspect. In the case of self-enhancement, specific factors might correspond to subdomains such as "narcissism" or "moral identity," capturing characteristics unique to these domains that are not fully explained by the general factor.

The bifactor model is a powerful statistical tool that allows for a deeper understanding of the internal structure of complex constructs. It enables the distinction between general trends and specific manifestations while also providing insights into the validity of measurement tools. This model has found widespread applications in psychology, education, medicine, and other fields. To assess the applicability of the model, model fit indices such as RMSEA, CFI, and TLI were reported.

3.3.4 Random Forest Analysis

The random forest algorithm is an ensemble learning method widely used for classification and regression tasks in machine learning. It improves overall predictive accuracy and robustness by constructing multiple decision trees and aggregating their predictions. In this study, random forest regression was utilized to predict real-world outcomes (e.g., depression, procrastination) based on principal components extracted from questionnaire and task data.

Feature importance metrics, including Mean Decrease Accuracy and Mean Decrease Gini, were employed to identify the most predictive variables. The model's performance was evaluated using Mean Absolute Error (MAE), Mean Squared Error (MSE), and R-squared (R²).

Discussion

Key Findings

Data analysis yielded ...... The findings of the study validated ...... the hypothesis of.

The ontology of Self-enhancement

Discuss how findings advance understanding of self-enhancement’s ontology.

Practical Applications

Self-enhancing predictive roles in mental health prevention and treatment.

Limitations

Future Directions

Ontological Studies.

Cross-cultural studies.

**References**

**[1] Judge TA. Erez A, Bono JE，et al.The core self-evaluations scale: Development of a measure.Personnel Psychology,2003，56(2)303-331**

**Appendix**