

MASTERARBEIT | MASTER'S THESIS

Titel | Title

Objective Personality Tests sensu Cattell: Test construction and
validation

verfasst von | submitted by
Ana-Maria Zerbes BSc

angestrebter akademischer Grad | in partial fulfilment of the requirements for the degree of
Master of Science (MSc)

Wien | Vienna, 2026

Studienkennzahl lt. Studienblatt | Degree
programme code as it appears on the
student record sheet:

UA 066 840

Studienrichtung lt. Studienblatt | Degree
programme as it appears on the student
record sheet:

Masterstudium Psychologie

Betreut von | Supervisor:

Assoz. Prof. Mag. Dr. Jakob Pietschnig Privatdoz.

Danksagung

Diese Arbeit widme ich allen, die mich auf diesem Weg begleitet haben:

Mein besonderer Dank gilt meinem Betreuer, Assoz. Prof. Mag. Dr. Jakob Pietschnig, für das Vertrauen, das er mir mit diesem Thema entgegengebracht hat. Ich danke ihm für seine kontinuierliche Unterstützung, seine wertvollen Rückmeldungen und die vielen hilfreichen Gespräche während der Arbeit. Besonders dankbar bin ich ihm für seine Ermutigung und seinen Glauben an meine Fähigkeiten, die mich während des gesamten Projekts und darüber hinaus motiviert und inspiriert haben.

Mein Dank gilt meiner Familie, insbesondere meiner Mutter, der ich vieles zu verdanken habe. Ihre Ermutigung und Unterstützung haben mich immer begleitet und wesentlich dazu beigetragen, dass ich diesen Weg gehen konnte. Gleichzeitig danke ich Andi, der diese Arbeit auf vielfältige Weise unterstützt hat, indem er mir zugehört hat und mir gerade in intensiven Phasen den nötigen Rückhalt gegeben hat.

Ich danke meinen Freundinnen Romina, Nadja und Mona für ihre Unterstützung, für das Durchgehen der Pretest-Items und für die Hilfe beim Austeilen der Flyer für meine Studie.

Danke auch an alle Studienteilnehmerinnen und Studienteilnehmer, die sich die Zeit genommen haben, an meiner Studie mitzumachen und damit nicht nur zu meinem Studienabschluss, sondern auch zum wissenschaftlichen Fortschritt beigetragen haben.

Schließlich danke ich dem Evangelischen Studienwerk Villigst herzlich für die langjährige Unterstützung. Durch diese Förderung haben sich für mich viele Möglichkeiten eröffnet, für die ich immer dankbar sein werde.

Ein letzter Dank gilt Emma, meiner „feline assistant“, die das Verfassen dieser Arbeit oft deutlich unterhaltsamer gemacht hat.

Table of Contents

Theoretical Background.....	5
Present study	13
<i>Approval-based Self-Worth</i>	13
<i>Principle-Based Self-Worth</i>	14
Method and Materials	17
Test Development	17
<i>Item Development for Approval-Based Self-Worth Test</i>	18
<i>Item Development for Principle-Based Self-Worth Test</i>	20
Procedure and Participants.....	22
Measures	23
<i>Newly Developed Objective Subtests</i>	23
<i>Contingencies of Self-Worth Scale</i>	24
<i>HEXACO Personality Inventory-Revised (HEXACO-PI-R)</i>	25
<i>International Cognitive Ability Resource (ICAR)</i>	25
<i>Original OA-TB 75 Subtests</i>	25
Analysis Plan	27
Results.....	28
Structural Analyses	28
Convergent and Discriminant Validity	32
Original OA-TB Subtests Structure	33
Discussion.....	33
Associations with Self-Report Measures and General Cognitive Ability	35
Original OA-TB Subtests.....	36
Implications	36
Limitations	37
Conclusion	38
References.....	39
List of Tables	49
List of Abbreviations	49
Appendix.....	50
Abstract (English)	50
Abstract (German)	51

Objective Personality Tests *sensu* Cattell: Test construction and validation

Contemporary assessment of personality traits relies predominantly on self-report questionnaires (Ashton & Lee, 2025; Kim et al., 2019; Phan & Rauthmann, 2021). These instruments are efficient, economical, and easy to administer, which led to their widespread adoption in research and applied contexts (Paulhus & Vazire, 2007; Pavlov et al., 2021). Moreover, they have demonstrated substantial predictive validity for a wide range of behaviors (Corneille & Gawronski, 2024; Pozzebon & Ashton, 2009). However, being based on introspection and deliberate self-presentation, they are also vulnerable to systematic response biases (Paulhus, 2017). Social desirability, impression management, and limited self-insight may distort responses, thereby potentially threatening the validity of the assessment (Latkin et al., 2017; Paulhus, 1984).

An alternative approach was proposed by Raymond B. Cattell through objective personality tests (Cattell, 1948), that aim to infer trait variance from behavioral responses to standardized tasks, without participants' explicit awareness of the trait being assessed. Despite its conceptual sophistication, objective personality testing *sensu* Cattell has received comparatively limited attention in contemporary psychometric research (Santacreu et al., 2006). Following their initial development, they were first revisited during the rise of computerized assessment in the 1990s, which enabled more efficient administration and scoring of behavioral tasks (Ortner & Schmitt, 2014). However, in many cases, objective tasks became increasingly detached from Cattell's source trait model and were integrated into assessment systems without explicit reference to their conceptual foundations. Importantly, Cattell did not conceive objective testing as a closed system but as an evolving research program. He explicitly encouraged further refinement, empirical examination, and expansion of objective measures to strengthen their theoretical and psychometric basis.

Moreover, theoretical integration between Cattell's objective factors and contemporary personality and motivational constructs remained underdeveloped. Two such objective factors—U.I. 29 (Determined Responsiveness vs. Lack of Will) and U.I. 36 (Strong vs. Weak Self-Sentiment; Cattell & Warburton, 1967)—can be conceptually reframed as reflecting distinct modes of self-worth regulation. Whereas one dimension may capture approval-oriented responsiveness and externally contingent self-evaluation, the other may reflect principle-based self-regulation, moral consistency, and internalized standards of conduct.

The present study aimed to develop and validate a set of newly constructed objective subtests designed to assess two theoretically distinct forms of self-worth regulation corresponding to U.I. 29 and U.I. 36 (Cattell & Warburton, 1967). Six behavioral subtests were constructed in a scenario-based questionnaire format inspired by the objective test logic underlying the Objective-Analytic Test Battery. Furthermore, the study psychometrically evaluated the internal structure, reliability, and their convergent and discriminant validity with contemporary personality and self-worth measures. By combining historical theoretical foundations with modern dimensions, this study aims to contribute to a renewed methodological discussion regarding objective personality assessment.

Theoretical Background

Personality traits are conceptualized as relatively stable behavioral dispositions within a person that manifest in consistent patterns of thinking, feeling, and acting across situations (Johnson, 1997). Since these traits are associated with a wide range of life outcomes (Roberts et al., 2007), including academic achievement (Chen et al., 2025; Mammadov, 2022), health (Strickhouser et al., 2017; Weston et al., 2015), and well-being (Anglim et al., 2020; Joshanloo, 2022), their accurate assessment is a central interest in psychological research and applied contexts (Pinon, 2019). In psychology, personality trait assessment relies predominantly on self-report questionnaires (Olinio & Klein, 2015). Although these instruments offer efficient, standardized, and economical data collection, they depend on introspection and deliberate self-description (Corneille & Gawronski, 2024; Paulhus & Vazire, 2007). As a result, responses may be influenced by response bias such as social desirability, as well as limited self-insight (McDonald, 2008; Paulhus, 1984).

Recognizing these methodological constraints more than half a century ago, Raymond B. Cattell proposed that personality should be assessed through multiple data sources to capture individual differences more comprehensively and with less bias (Cattell, 1948). He distinguished life record data (L-data) from questionnaire data (Q-data) and objective test data (T-data), arguing that each source provides complementary information about underlying personality traits (Cattell, 1958).

While L-data refers to observations of behavior in everyday life (e.g., through behavioral ratings by observers) or recordings of life events (e.g., achievements), its assessment is reliable only under very specific research conditions, as it is susceptible to observer bias

(Cattell & Warburton, 1967). Similarly, Q-data, which relies on introspection and self-evaluations obtained through self-report questionnaires, can be subject to self-deception and faking (Cattell, 1958). Moreover, even when combined, L-data and Q-data do not cover the entire sphere of personality (Cattell & Warburton, 1967). T-data was intended to overcome these limitations by providing the highest level of objectivity through standardized test situations and scoring rules. Most importantly, behavior is assessed without the test taker knowing what trait is being measured, with the aim of reducing systematic bias (Cattell & Warburton, 1967).

In contrast to ability assessment, where performance based objective measures had long been established, comparable objective approaches to personality traits were largely underdeveloped (Cattell, 1955). To address this gap, Cattell introduced the concept of *Objective Personality Tests* (OPTs; Cattell, 1948). OPTs were defined as standardized behavioral assessment situations in which the purpose of measurement is not transparent to the person taking the test, with responses scored according to objective rules, and trait-relevant reactions elicited through controlled task conditions (Cattell & Warburton, 1967). The goal is, on the one hand, to reduce the extent to which respondents can intentionally distort their answers in accordance with perceived social desirability. On the other hand, OPTs aim to bypass potential limitations of introspection, recognizing that individuals may lack full insight into their own dispositional tendencies (Cattell, 1958).

In their compendium, Cattell and Warburton (1967) presented a little over 600 OPTs and over 2,400 derived variables, which constituted the actual operational unit of measurement within the test (e.g., completion time, number of errors, or convergence indices between responses). Initial tests were developed through experimental approaches and rooted in factor analysis of pre-existing clinical and questionnaire data. Subsequently, standardized situations were designed to elicit behavior hypothesized to be related to a specific trait (Cattell & Warburton, 1967).

Through extensive factor-analytic investigations across large sets objective tests, Cattell and his collaborators identified 21 source traits, labeled U.I. 16 to U.I. 36. The labels were part of a taxonomy that aimed to provide a structured and cumulative system for researchers with the first 15 factors representing ability traits. This hierarchical indexing system contained extracted source traits (i.e., factors, labeled U.I. for Universal Index), the individual objective

tests (T), and the behavioral variables derived from those tests (Master Index, M.I.) (Cattell & Warburton, 1967).

Although Cattell explicitly envisioned that subsequent research would refine, expand, and validate the tests presented in the compendium, comparatively little systematic follow-up work occurred, one of them being a German version of 50 tests, the *Objective-Analytic Test Battery* (OA-TB) by Häcker et al. (1975). Renewed interest in objective personality assessment emerged in the 1990s alongside the increasing digitalization of psychological testing.

Computerized administration enabled more precise behavioral recording and facilitated experimental task designs (Ortner & Proyer, 2015). However, this second generation of OPTs was only loosely grounded in Cattell's original framework, while conceptually, the focus shifted toward what was termed experiment-based assessment of behavior (Kubinger, 2009; Ortner & Schmitt, 2014). Moreover, modern OPTs have focused on a limited number of constructs such as risk-taking propensity, conscientiousness, and frustration tolerance. Prominent examples include the *Balloon Analogue Risk Task* (Lejuez et al., 2002), the *Vienna Risk-Taking Test* (Hergovich et al., 2007), *Working Styles* (Kubinger & Ebenhö, 1996; 2002) and objective task-based measures of conscientiousness (e.g., Koch et al., 2014). These instruments assess trait-relevant tendencies through standardized behavioral tasks (e.g., simulated situations, achievement tasks) without it being transparent which construct is being measured, therefore sharing the general logic of objective assessment originally proposed by Cattell and Warburton (1967).

Despite this conceptual overlap, most of them are not based on the original tests from the compendium, nor are they grounded in the behavioral descriptions and trait definitions proposed by Cattell. Moreover, systematic validation and cumulative development of the original OPTs have remained limited. Much of the early OPT material has not been re-examined, and direct conceptual continuations of this assessment approach are scarce (Ortner & Schmitt, 2014; Romero et al., 2024). As a result, the historical line of research on objective personality measurement has become fragmented, with modern assessment evolving largely independently from the original framework.

This discontinuity suggests a gap in psychometric personality research. Revisiting objective assessment combined with Cattell's trait conceptualization may allow a clearer picture on

how OPTs can be developed and how they perform today. In particular, paper-and-pencil or questionnaire-like behavioral formats conceptually aligned with Cattell's approach may not only provide a theoretically grounded pathway for renewed objective test construction but also increase test accessibility, as modern OPTs rely on specific software and technical requirements (Romero et al., 2024).

When examining the 21 objective source traits described by Cattell and Warburton (1967) in the compendium, several factors can be readily related to constructs that have recognizable counterparts in contemporary personality psychology (e.g., extraversion-introversion or reflexivity-impulsivity). These factors emerged empirically from extensive factor analyses of behavioral variables. Consequently, the description of each source trait is grounded in the pattern of variables that loaded on the respective factor, rather than in verbal trait descriptions alone (Cattell & Warburton, 1967). Within this taxonomy, however, the empirical robustness of factors varied. Some source traits were supported by multiple tests and a relatively coherent set of variables, whereas others showed weaker structural definition and required further empirical elaboration (Cattell & Warburton, 1967; Häcker et al., 1975). One such factor is U.I. 36, for which Cattell noted the need for additional research and test development to clarify its structure. Thematically related to U.I. 36 is its U.I. 29, which appeared somewhat more stable in the analyses and was represented by a broader set of variables (Cattell & Warburton, 1967). This asymmetry is also reflected in the German adaptation of the Objective-Analytic Test Battery (Häcker et al., 1975), where three tests were assigned to U.I. 29 but only a single test to U.I. 36. The comparatively limited operationalization of U.I. 36 suggests that this factor remained less fully developed within the original objective assessment framework. As a result, both factors—particularly U.I. 36—represent areas in which further objective test construction and validation appear warranted.

To relate these factors to contemporary personality constructs, the present study draws on the original behavioral descriptions provided in the compendium (Cattell & Warburton, 1967). On this basis, U.I. 29 and U.I. 36 can be interpreted as reflecting distinct forms of self-regulatory orientation, differing in the extent to which behavior is guided by external approval versus internalized standards. The following section therefore outlines the conceptual translation of these factors into modern theoretical terms.

U.I. 29, labeled *Wholehearted Responsiveness (vs. Lack of Will)*, describes a pattern of heightened reactivity to immediate social demands combined with a pronounced readiness to engage in the present situation (Cattell & Warburton, 1967). Individuals scoring high on this factor are characterized by a tendency toward heightened emotional reactivity to situational cues, yet they also relax rapidly once the demand subsides. Their behavior in structured settings suggests alertness to interpersonal signals, particularly those conveyed by authority figures or evaluators. A central element of this factor is a strong desire to please others, especially within evaluative contexts. High scorers respond positively to approval and show measurable sensitivity to disapproval, although they may display relatively less impairment under disapproval compared to individuals low on the factor. As outlined by Cattell and Warburton (1967) their performance often improves under conditions of approval, reflecting a motivational orientation toward social reinforcement. They may show immediate efficiency, flexibility, and good short-term memory under structured and supervised conditions. However, this rapid engagement is accompanied by tendencies toward overresponsiveness. High scorers may exhibit increased proneness to false reactions, decreased accuracy when not regularly prompted, and reduced steadiness in tasks requiring sustained, solitary effort. This pattern has been associated with negative correlations with scholastic and technical training performance, suggesting difficulty maintaining consistent performance in the absence of external structure or feedback (Cattell & Warburton, 1967). At the cognitive-attitudinal level, U.I. 29 is positively associated with the acceptance of conventional aphorisms. Individuals high on the factor are more likely to accept socially endorsed statements without critical examination, reflecting a readiness to align with prevailing norms and expressed opinions. Conversely, critical or subtly hostile evaluative tendencies show negative associations with the factor, indicating that low scorers may display greater skepticism, defensiveness, or resistant criticalness (Cattell & Warburton, 1967).

Overall, U.I. 29 aimed to reflect a temperament characterized by rapid situational engagement, social responsiveness, and motivational sensitivity to approval cues, combined with reduced resistance to conventional ideas and potential difficulty sustaining independent, self-directed effort.

As described by Cattell and Warburton (1967), the second factor, U.I. 36, *Strong Self Sentiment* (vs. *Low Self Sentiment*), is characterized by strong loadings on measures of control and self-respect and represents one of the clearest objective indicators of the dynamic self-sentiment factor within Cattell's framework. Among the objective factors, it shows particularly substantial associations with second-order ego strength and alertness, suggesting that it reflects a fundamental dimension of organized self-regulation (Cattell & Warburton, 1967).

Individuals scoring high on U.I. 36 exhibit a pattern of controlled, alert, and assured behavior. They demonstrate wariness and low levels of impulsivity. As described by Cattell and Warburton (1967), their behavioral profile suggests stable inhibition and disciplined regulation of action, accompanied by relatively low anxiety. At the motivational level, high scorers are described as possessing a strong desire to maintain control over their mental processes and thinking. They show reluctance to engage in behaviors that might threaten or diminish their self-respect. This controlled self-regard appears to function as an organizing principle of behavior, shaping decision-making in situations that involve evaluative or potentially self-threatening elements (Cattell & Warburton, 1967).

Physiological indicators further support this characterization. High scorers demonstrate heightened psychogalvanic reflex responses to threatening stimuli, suggesting underlying fearfulness or sensitivity to threat. At the same time, they exhibit less upward physiological drift during relaxation compared to shock conditions, indicating sustained tension and vigilance. These findings suggest that their behavioral control may be accompanied by a cautious or watchful orientation toward potential risks (Cattell & Warburton, 1967). Overall, U.I. 36 represents a pattern of alert but assured self-control, organized around the maintenance of self-respect and disciplined regulation of thought and action.

Although conceptually related, the two factors therefore represent distinct regulatory orientations: one organized around social evaluation and approval (U.I. 29), and the other around internalized principles and self-consistency (U.I. 36). The distinction between externally guided and internally guided regulation has also been emphasized in broader theories of behavioral control. External regulation refers to behavior guided by anticipated evaluation, approval, or disapproval from others, whereas internal regulation reflects behavior aligned with personally endorsed values and standards (Howard et al., 2021).

Individuals differ in the extent to which they orient behavior toward external versus internal criteria, implying different motivational bases of action. Behavioral regulation is closely linked to the regulation of self-worth: When individuals' sense of personal value depends on success or validation in specific domains, behavior becomes oriented toward protecting self-worth (Crocker et al., 2006). Such ego-involved or introjected forms of regulation can increase effort and performance when success appears attainable, but they may also lead to avoidance, anxiety, and reduced intrinsic motivation when failure is anticipated. In these cases, the protection of self-worth can override optimal task engagement, and behavior may be shaped more strongly by the avoidance of shame or negative evaluation than by effective goal pursuit (Crocker et al., 2006).

The *contingencies-of-self-worth model* (CSW; Crocker & Wolfe, 2001) proposes that individuals differ in the domains on which their sense of personal value depends. Although related to global self-esteem, contingencies of self-worth refer not to overall self-evaluation, but to the domains and conditions upon which self-worth depends. These contingencies can be externally grounded, such as approval from others, or internally grounded, such as adherence to personal moral standards or virtue (Crocker & Wolfe, 2001). When self-worth is contingent on external approval, behavior may become oriented toward obtaining positive evaluation and avoiding disapproval. In contrast, when self-worth is grounded in internal standards, behavior may be guided more strongly by personal values and the maintenance of self-consistency.

Within the CSW framework (Crocker & Wolfe, 2001), U.I. 29 may be interpreted as reflecting a stronger orientation toward externally contingent self-worth, whereas U.I. 36 appears consistent with regulation grounded in internally endorsed standards and controlled self-respect. Although the constructs are not identical, the conceptual overlap suggests that Cattell's objective factors may capture behavioral expressions of distinct self-worth contingencies.

The distinction between externally contingent and internally grounded self-worth regulation also resonates with motivational differentiation within *self-determination theory* (SDT; Deci & Ryan, 1985). SDT conceptualizes behavior as varying in the degree to which it is regulated by external pressures versus internalized standards. Introjected regulation refers to behavior driven by internal pressure to avoid guilt, shame, or loss of self-worth, whereas more

autonomous forms of regulation (identified or integrated regulation) involve behavior aligned with personally endorsed values (Ryan & Deci, 2020).

The orientation reflected in U.I. 29, characterized by heightened sensitivity to evaluation and approval-oriented responding (Cattell & Warburton, 1967), shows conceptual similarity to introjected forms of regulation in which behavior is shaped by concerns about external validation and self-worth contingencies. U.I. 36, in contrast, reflects regulation organized around the maintenance of internally defined standards and controlled self-respect. While not identical to autonomous regulation as defined in SDT, this orientation shares the emphasis on internally referenced standards rather than immediate social evaluation.

Together, these theoretical parallels suggest that the behavioral tendencies captured by the two objective factors may reflect distinct motivational orientations underlying self-worth regulation. The objective factors U.I. 29 and U.I. 36 (Cattell & Warburton, 1967) describe these orientations at a behavioral level, Crocker and Wolfe's (2001) contingencies-of-self-worth framework conceptualizes them as distinct bases of self-evaluation, and self-determination theory (Deci & Ryan, 1985) characterizes them as differing motivational regulation modes. Across these traditions, a consistent distinction emerges between externally contingent and internally grounded regulation of behavior and self-worth.

Traditional assessment of self-esteem has relied predominantly on explicit self-report measures (Donnellan et al., 2015) such as the Rosenberg Self-Esteem Scale (Rosenberg, 1965). While widely validated, such instruments measure perceived self-worth rather than the regulatory processes through which self-worth is maintained.

Attempts to assess implicit self-esteem using paradigms such as the Implicit Association Test (IAT; Greenwald et al., 1998) seek to reduce deliberate self-presentation. However, these measures capture associative evaluations rather than stable behavioral regulatory tendencies and have demonstrated mixed reliability and validity across contexts (Greenwald & Farnham, 2000; Jusepeitis & Rothermund, 2022; Krause et al., 2011; Schimmack & Diener, 2003).

Importantly, neither explicit self-report measures nor implicit association paradigms directly assess how individuals regulate behavior under conditions that threaten or affirm self-worth. The behavioral expression of self-worth contingencies—particularly within standardized task situations—remains therefore largely unexamined. Specifically, no objective personality test has been developed to measure contingencies of self-worth within a behavior-based trait

framework. This represents a notable gap, particularly in light of Cattell and Warburton's (1967) original proposal that stable regulatory tendencies should be measured through standardized behavioral situations rather than inferred solely from self-descriptions.

Present study

The present thesis addresses this gap by developing and examining an objective personality test grounded in Cattell's objective assessment tradition. Building on the methodological principles and theoretical descriptions laid out by Cattell and Warburton (1967) and further systematized in the German Objective-Analytic Test Battery (Häcker et al., 1975), the present study develops new objective subtests designed to operationalize behavioral expressions of externally versus internally contingent self-worth. Specifically, the Cattellian factors U.I. 29 and U.I. 36 are translated into contemporary theoretical terms and operationalized as approval-based and principle-based self-worth regulation. Although the subtests are administered in questionnaire format, they follow Cattell's objective assessment principle by presenting structured third-person material and deriving trait inferences from patterned response tendencies rather than from direct self-evaluation. This approach aligns with the T-data tradition in which personality is inferred from performance and decision tendencies within standardized situations (Cattell & Warburton, 1967). The primary aim is to examine the psychometric properties of these newly developed objective subtests, including their internal structure, reliability, and convergent and discriminant associations with established self-report measures of contingencies of self-worth and related constructs. To operationalize approval-based and principle-based self-worth regulation in objective behavioral terms, the present test development focuses on specific behavioral dimensions derived from Cattell's factor descriptions and their contemporary interpretation. Each dimension represents a distinct way in which externally or internally contingent self-worth may be expressed in behavior. The following sections outline the conceptual basis of each dimension as described by Cattell (Cattell & Warburton, 1967).

Approval-based Self-Worth

Dimension 1: Sensitivity to Approval and Disapproval. Sensitivity to social evaluation reflects the tendency to monitor and respond to cues of approval or disapproval in interpersonal contexts. In Cattell and Warburton's (1967) description of U.I. 29, individuals high on this factor react positively to approval and show heightened responsiveness to social

feedback. They are alert to social demands and may quickly adjust their behavior to appear engaged or competent. Behaviorally, this pattern may manifest in frequent checking for feedback, rapid adjustment of expressed positions in group settings, and heightened responsiveness to evaluative cues (Cattell & Warburton, 1967). The corresponding subtest presents social situations in which response options vary in the degree to which they signal monitoring of and adaptation to social evaluation. Higher scores indicate stronger orientation toward externally anchored regulation of behavior.

Dimension 2: Desire to Please Others. Desire to please and acceptance of socially endorsed positions reflects a readiness to align with the ideas and expectations of others in order to maintain interpersonal harmony and positive evaluation. Individuals high in U.I. 29 display greater acceptance of conventional statements and less critical evaluation of others. Behaviorally, this orientation may be expressed through agreement with majority opinions, endorsement of conventional viewpoints, and willingness to engage in tasks primarily to satisfy external expectations (Cattell & Warburton, 1967). The corresponding subtest captures this tendency through scenarios in which response options differ in the degree of social conformity versus independent positioning. Higher scores indicate preference for socially aligned responses over critical or autonomous alternatives.

Dimension 3: Poor Solitary Performance. This dimension reflects the tendency to perform more efficiently and with greater engagement when behavior is visible to others. Cattell and Warburton (1967) described individuals high in U.I. 29 as showing increased efficiency under immediate social demands but reduced steadiness in solitary contexts. Behaviorally, this may manifest as preference for tasks that are observable by others, increased speed or engagement when being watched, and reduced motivation in private, non-evaluative settings. The corresponding subtest presents work-related scenarios in which response options vary in the degree of visibility and evaluative salience. Higher scores indicate stronger reliance on external observation as a motivational regulator.

Principle-Based Self-Worth

Dimension 1: Desire for Mental Self-Control. Desire for mental self-control reflects the tendency to regulate one's thoughts and actions in accordance with internally defined standards of discipline and self-respect. In Cattell and Warburton's (1967) conceptualization, individuals high in U.I. 36 show alert but assured self-control, steady

inhibition, and low impulsivity. They display reluctance to engage in behavior that might undermine their self-respect and are motivated to maintain control over their mental processes. Behaviorally, this may manifest in preference for structured routines, adherence to long-term goals, and avoidance of situations in which one's self-image could be compromised. The corresponding subtest presents scenarios in which response options differ in the degree of deliberate self-regulation versus spontaneous or convenience-based responding. Higher scores indicate stronger orientation toward controlled and internally regulated action.

Dimension 2: Sensitivity to Mistakes. Sensitivity to mistakes and vigilant task execution reflects heightened concern about committing errors that could threaten one's self-respect. Cattell and Warburton (1967) portrayed U.I. 36 as associated with increased psychogalvanic responses to threat, suggesting underlying vigilance and cautiousness. Individuals high on this factor exhibit behavioral control and low impulsivity, and this control may be accompanied by or rooted in carefulness or even heightened concern about potential mistakes. Behaviorally, this dimension may manifest in thorough checking of work, careful execution of instructions, and preference for accuracy over speed. The corresponding subtest captures decision patterns that reflect cautious versus expedient task engagement. Higher scores indicate stronger orientation toward careful, self-protective regulation of performance.

Dimension 3: Consistency Between Behavior and Personal Principles. Consistency between behavior and personal principles reflects the tendency to prioritize internally endorsed values over situational pressures or social trends. Within Cattell and Warburton's (1967) framework, U.I. 36 represents controlled self-respect grounded in adherence to personal standards. Individuals high on this dimension are expected to align their behavior with internal principles, even when such choices involve inconvenience or potential social costs. Behaviorally, this may be expressed in selecting value-consistent actions over socially convenient alternatives, resisting conformity pressures, and maintaining position in the face of external influence. The corresponding subtest presents scenarios in which response options vary in the degree to which they prioritize personal values versus social approval or pragmatic ease. Higher scores indicate stronger principle-based regulation.

Based on the theoretical distinction between approval-based and principle-based self-worth regulation, the newly developed objective subtests are expected to reflect two coherent and distinguishable regulatory orientations. Specifically, the six subtests are expected to cluster into two higher-order dimensions corresponding to approval-based self-worth regulation (U.I. 29) and principle-based self-worth regulation (U.I. 36) (H1). Furthermore, a two-dimensional representation of the subtests is expected to provide a better account of their interrelations than a unidimensional structure (H2). Although conceptually related, the two regulatory orientations are assumed to represent distinct modes of behavioral regulation; therefore, the two dimensions are expected to be positively correlated but clearly distinguishable (H3). Consistent with the assumption that they capture stable regulatory tendencies, each subtest and higher-order dimension is expected to demonstrate adequate internal consistency (H4).

Beyond internal structure and reliability, the construct validity of the two regulatory orientations will be examined through their associations with established measures of self-worth contingencies and personality traits. In line with contingencies-of-self-worth theory, approval-based self-worth regulation is expected to correlate positively with the Approval from Others subscale of the Contingencies of Self-Worth Scale (Crocker et al., 2003), reflecting externally contingent self-worth (H5), whereas principle-based self-worth regulation is expected to correlate positively with the Virtue subscale (Crocker et al., 2003), reflecting self-worth grounded in internalized moral standards (H6).

In addition to contingencies-of-self-worth model, the proposed regulatory orientations can be situated within broader personality trait models. The HEXACO model of personality (Ashton & Lee, 2008; 2017) provides a comprehensive framework for examining individual differences in honesty-humility, emotionality, extraversion, agreeableness, conscientiousness, and openness. Because approval-based and principle-based regulation forms reflect distinct motivational and evaluative tendencies, examining their associations with selected HEXACO facets allow the new objective measures to be located within the broader nomological network of personality traits. Distinct patterns of association are further expected with HEXACO personality facets (Ashton & Lee, 2008). Specifically, approval-based self-worth regulation is expected to correlate positively with Gentleness and Flexibility facets (Agreeableness dimension) as well as Modesty (Honesty-Humility dimension),

reflecting socially attuned and approval-sensitive tendencies (H7). In contrast, principle-based self-worth regulation is expected to correlate positively with Sincerity (Honesty-Humility dimension), as well as Perfectionism and Prudence (Conscientiousness dimension), reflecting principled and self-controlled behavioral regulation (H8).

To establish discriminant validity, associations with cognitive ability will also be examined. Because the proposed measures are intended to capture personality-related regulatory processes rather than intellectual performance, both regulatory dimensions are expected to show trivial correlations with the short form of the International Cognitive Ability Resource (ICAR; Condon & Revelle, 2014) (H9).

Finally, the study explores the correspondence between the newly developed measures and selected original OA-TB 75 tests. With regard to their conceptual associations, the original OA-TB tests linked to U.I. 29 (T8, T18, T111) are expected to cluster together, whereas the test associated with U.I. 36 (T9) is expected to show low or non-salient association with this grouping (H10). In line with the factor loadings reported by Cattell and Warburton (1967), approval-based self-worth regulation is expected to correlate negatively with T8 and T111 and positively with T18 (H11), and principle-based self-worth regulation is expected to correlate negatively with T9 (H12).

Method and Materials

Test Development

The newly developed objective subtests were constructed based on the methodological principles and theoretical descriptions articulated by Cattell and Warburton (1967), and the German version of the Objective-Analytic Test Battery (OA-TB 75; Häcker et al., 1975). These sources provided both conceptual guidance regarding the behavioral expression of U.I. 29 and U.I. 36 and practical inspiration for task format and structure. In line with the objective assessment tradition presented by Cattell and Warburton (1967), the present measures were designed in a standardized questionnaire format. This format was chosen to ensure ease of administration, replicability, and applicability in both research and applied contexts. Moreover, many of the original objective tests described by Cattell and Warburton (1967) were implemented in questionnaire format, including scenario evaluations, continuation of short narratives, and forced-choice judgments. Adopting a comparable format

therefore allowed the present test development to remain consistent with the original methodological approach.

The development process followed a theory-driven structure. For each factor (U.I. 29 and U.I. 36), the three distinct dimensions were operationalized through a dedicated subtest. Items were formulated in third-person, scenario-based formats to elicit patterned response tendencies without directly asking participants to describe themselves. Consistent with the recommendation to minimize response fatigue (Cattell & Warburton, 1967), multiple objective formats were employed, including situational evaluations and behavioral judgments. This variation in format aligns with the recommendation to diversify objective task modes in order to reduce response sets and enhance construct coverage (Cattell & Warburton, 1967). Higher scores reflect stronger endorsement of externally contingent (U.I. 29) or internally grounded (U.I. 36) regulatory orientations, respectively.

Item Development for Approval-Based Self-Worth Test

Dimension 1: Sensitivity to Approval and Disapproval. The first subtest, *Feedback Sensitivity* (SF) operationalizes sensitivity to social evaluation as a core component of approval-based self-worth regulation. In line with Cattell and Warburton's (1967) description of U.I. 29 characterized by heightened responsiveness to approval and disapproval, the subtest was designed to capture individual differences in the interpretation of ambiguous social cues as evaluative and self-relevant. Items consist of brief, everyday social situations involving subtle interpersonal signals, such as nonverbal reactions, topic changes, indirect feedback, or comparative praise. Each scenario is followed by four response options that systematically vary in the degree to which they reflect (a) strong self-referential interpretation of the situation as evaluative, (b) moderately socially anchored interpretation, (c) neutral continuation without evaluative inference, or (d) attribution to situational or non-self-related causes. This graded structure allows for indirect assessment of the extent to which individuals spontaneously interpret social cues as meaningful indicators of approval or disapproval. All items were formulated in third-person format to avoid direct self-description and to ensure that trait inferences are derived from patterned response tendencies rather than explicit self-evaluations. Higher scores reflect a stronger tendency to interpret ambiguous interpersonal situations as evaluative and self-relevant, consistent with externally contingent regulation of self-worth.

Dimension 2: Desire to Please Others. The second subtest, *Conformity Appraisal* (CA) operationalizes the tendency to endorse or normalize behavior aimed at maintaining interpersonal approval. Consistent with the description of U.I. 29 as reflecting a wholehearted readiness to accept the ideas of others and a reduced inclination toward critical opposition (Cattell & Warburton, 1967), this dimension captures the degree to which approval-oriented behaviors are perceived as understandable or appropriate. Items describe everyday interpersonal behaviors that involve aligning with others, avoiding disagreement, adapting one's position, or prioritizing interpersonal harmony over independent assertion. Rather than asking participants to report their own behavior, the subtest employs an evaluative format in which respondents rate how understandable such behaviors appear to them on a four-point scale ranging from 1 (*not understandable at all*) to 4 (*very understandable*). This format allows indirect assessment of approval-oriented regulation by measuring the extent to which behaviors aimed at pleasing others or avoiding interpersonal friction are perceived as normatively justified. Higher ratings indicate stronger acceptance of socially accommodating and conformity-oriented behaviors, consistent with externally contingent self-worth regulation. Lower ratings reflect greater distance from such appeasing tendencies and stronger endorsement of independent positioning.

Dimension 3: Poor Solitary Performance. The third subtest, *Task Preference* (TP) operationalizes audience-dependent engagement, reflecting the tendency to prefer task contexts in which performance is socially visible over contexts in which it remains private. This dimension is grounded in Cattell and Warburton's (1967) observation that individuals high in U.I. 29 tend to show increased engagement and responsiveness in socially present or evaluative situations, whereas solitary, non-observed work may be less motivating. Each item presents two structurally comparable activities (A and B). Option A represents an independently conducted task without interpersonal visibility or feedback, whereas Option B represents a socially embedded task context characterized by collaboration, shared progress, or regular exchange with others. Participants indicate their typical preference on a four-point scale ranging from 1 (*clearly A*) to 4 (*clearly B*), thereby allowing graded expression of directional preference. The absence of a neutral midpoint was intended to encourage differentiation between socially embedded and solitary engagement tendencies. Importantly, the distinction between collaborative and solitary contexts also implicitly varies the

availability of interpersonal feedback and evaluative visibility. Situations involving joint work, exchange, or shared progress inherently increase exposure to observation and potential evaluation, whereas solitary contexts reduce such exposure. Rather than explicitly referencing approval or feedback, evaluative conditions are thus operationalized through differences in task structure. This approach is consistent with the objective testing logic, in which motivationally relevant features—such as competition or evaluation—are embedded in comparable task alternatives rather than stated directly (Cattell & Warburton, 1967). By varying social visibility without overt reference to approval, the subtest allows regulatory orientation toward evaluative environments to be inferred from consistent behavioral preferences. Higher scores (preference for socially embedded contexts) indicate stronger orientation toward externally visible and socially responsive performance situations, consistent with approval-based self-worth regulation. Lower scores reflect greater preference for independent, self-contained task engagement.

Item Development for Principle-Based Self-Worth Test

Dimension 1: Desire for Mental Self-Control. The first subtest for U.I. 36 operationalizes the desire for *structured mental self-control* (SR). In line with the description of U.I. 36 as reflecting controlled self-respect, inhibition, and deliberate regulation of thought processes (Cattell & Warburton, 1967), this dimension captures differences in how individuals approach demanding or goal-directed situations. Items describe everyday scenarios involving spontaneous demands, time pressure, distraction, or competing priorities. Each situation is followed by four possible reactions that vary systematically in the degree of structure, deliberation, and impulse control they reflect. Higher-scoring responses represent structured planning, prioritization, and completion-oriented behavior, whereas lower-scoring responses reflect spontaneity, postponement, or loosely organized action. All items are formulated in third-person format and require respondents to select the reaction that appears most fitting. By presenting behaviorally distinct alternatives rather than direct self-descriptions, the subtest allows regulatory orientation to be inferred from consistent preference patterns across situations. The response options were constructed to avoid obvious desirability cues; each alternative represents a plausible everyday reaction differing primarily in degree of mental organization and self-regulatory control. Higher total scores indicate a stronger tendency toward deliberate, structured, and inhibition-based regulation, consistent

with internally grounded self-worth protection. Lower scores reflect greater spontaneity, flexible responding, or reduced emphasis on structured control.

Dimension 2: Mistake Sensitivity. The second subtest, *Mistake Sensitivity* (MS), operationalizes vigilant carefulness, reflecting heightened attention to potential errors and preference for controlled task execution. This dimension is derived from Cattell and Warburton's (1967) characterization of U.I. 36 as involving reluctance to act in ways that could damage self-respect, as well as elevated vigilance and inhibition in performance contexts. Items describe everyday behaviors involving preparation, checking, correction, and adherence to instructions. Rather than asking participants to indicate how they would behave, the subtest presents a behavioral vignette followed by four trait descriptors: carefulness, caution, excessive control, and insecurity. Participants select the version that best characterizes the behavior in their opinion. This format allows indirect assessment of regulatory orientation by capturing how respondents interpret vigilant and error-preventive behavior. The response options represent gradations in perceived control orientation, ranging from overcontrol or insecurity to structured and deliberate carefulness. Higher scores reflect stronger endorsement of vigilant and control-oriented interpretations of behavior, consistent with a tendency to safeguard self-respect through careful execution and monitoring. Lower scores reflect reduced emphasis on controlled oversight or greater attribution to insecurity rather than principled regulation.

Dimension 3: Consistency Between Behavior and Personal Principles. The third subtest, *Principles* (PR), operationalizes principle-based consistency, reflecting the tendency to align behavior with internal standards even when competing considerations such as convenience, social harmony, or efficiency are present. This dimension captures the motivational aspect of U.I. 36 described as controlled self-respect and reluctance to engage in actions that could compromise personal standards (Cattell & Warburton, 1967). Each item presents a situation in which multiple considerations are simultaneously relevant (e.g., time pressure, group dynamics, effort minimization, or conflict avoidance). Four response alternatives represent distinct justificatory orientations. The highest-scoring alternatives reflect adherence to prior commitments, objective standards, or principled reasoning, even when this entails discomfort or increased effort. Intermediate alternatives reflect pragmatic compromise or balanced consideration of competing factors. Lower-scoring alternatives

prioritize convenience, impression management, or avoidance of interpersonal friction over principled consistency. The format requires respondents to select the response option that appears most fitting in general. By embedding value conflicts within everyday decision contexts rather than explicitly referencing moral identity, the subtest allows internalized standards to be derived from consistent choice tendencies. Higher aggregate scores indicate stronger preference for principled and standard-based justification of actions, consistent with internally grounded self-worth regulation. Lower scores reflect greater orientation toward convenience, social conformity, or situational expediency. This dimension complements the preceding subtests by extending mental self-control and vigilant carefulness into the domain of value-consistent action under social or practical pressure.

Procedure and Participants

Data collection took place online between October 27 and December 22, 2025, using the SoSci Survey platform. Participants were recruited through convenience sampling via university mailing lists, psychology-related social media groups (e.g., WhatsApp groups, Studo), brief announcements during psychology lectures at the University of Vienna, and distribution of flyers in lecture halls. Recruitment materials described the study as the validation of newly developed psychological questionnaires on decision making, without disclosing specific hypotheses or test content. Participants were required to be at least 18 years of age and to have sufficient proficiency in the German language to complete the survey. Participation was voluntary and anonymous. Before starting the questionnaire, participants received detailed study information and provided informed consent electronically. The estimated completion time was 50 minutes. At the end of the survey, participants were given the opportunity to enter a raffle for one of five €25 “Wunschgutschein” gift cards, with email addresses collected separately in an unlinked database to ensure anonymity of questionnaire responses. The planned recruitment target was approximately 240 participants, with the aim of obtaining around 200 valid cases after exclusions due to incomplete data or early termination. This target was based on expected dropout rates in online surveys of comparable length (Hoerger, 2010) and on psychometric recommendations suggesting that sample sizes of 150–200 participants are generally adequate for stable factor recovery in models of moderate complexity (Kyriazos, 2018; MacCallum et al., 1999). An a priori power analysis conducted using G*Power version

3.1.9.7 (Faul et al., 2007; 2009) for two-tailed correlations ($r = .20$, $\alpha = .05$, power = .80) further indicated that approximately 193 valid cases would be required to detect small-to-medium effects with sufficient statistical power.

A total of $N = 139$ individuals took part in the study. Missing data were examined descriptively. For the newly developed subtests, cases with more than 20% missing responses within any subtest (i.e., more than one missing item per seven-item subtest) were excluded from analyses, as reliable computation of subtest scores was not possible. For cases with minor missingness ($\leq 20\%$ of items within a subtest), subtest scores were calculated based on the individual's mean on the remaining items of the respective subtest. After applying these criteria, the final analytic sample comprised $N = 90$ participants. One further case was excluded due to implausible demographic information (reported age of 5 years) and an extreme response pattern characterized by selecting only minimum or maximum response options across measures.

The final sample ($N = 89$) had a mean age of $M = 26.56$ years ($SD_{age} = 7.24$, range = 20–65). The majority identified as female (79%) and residing in Austria (58%). Furthermore, 67% of participants were students at the time of participation. Across all participants, 58% reported holding a university or university of applied sciences degree. Completion time ranged from 5 to 64 minutes ($M_{time} = 38.64$, $SD_{time} = 14.43$). Due to the lower range, cases with completion time under 30 minutes ($N = 21$) were inspected for missing data and response patterns. Most of these cases ($N = 17$) corresponded to participants who dropped out before completing the entire survey. However, all of them provided responses to the developed subtests, which constituted the main measures of interest in the present study, and were therefore retained. Additionally, four participants completed the survey in less than 30 minutes but showed no unusual response patterns or reduced response variability and were also retained.

Measures

Newly Developed Objective Subtests

All subtests used four response options reflecting graded differences in regulatory orientation. Raw item responses were coded on a 1–4 scale, with higher values indicating stronger endorsement of the respective regulatory tendency. Response option order was randomized across items to prevent positional cues from signaling the direction of scoring. For each subtest, a mean score across the seven items was computed. The three subtest

means corresponding to each factor were subsequently averaged to form a composite score for approval-based self-worth regulation and principle-based self-worth regulation, respectively. Internal consistency of the six subtests was examined using Cronbach's α (Cronbach, 1951) and McDonald's ω (McDonald, 1999). Results are presented in Table 1:

Table 1

Internal Consistency Coefficients for the Developed (Sub)Tests

Variable	McDonald's ω	Cronbach's α
Approval-Based Self-Worth Regulation Test	.32	-.09 CI [-.51, .20]
Feedback Sensitivity (SF) Subtest	.33	.33 CI [.07, .50]
Conformity Appraisal (CA) Subtest	.78	.78 CI [.70, .83]
Task Preference (TP) Subtest	.56	.53 CI [.29, .68]
Principle-Based Self-Worth Regulation Test	.50	.21 CI [-.18, .46]
Structured Mental Self-Control (SR) Subtest	.55	.55 CI [.38, .67]
Mistake Sensitivity (MS) Subtest	.31	.14 CI [-.27, .39]
Principles (PR) Subtest	.44	.20 CI [-.04, .49]

Note. Confidence intervals are reported for Cronbach's α only (bootstrapped, 5,000 iterations). Confidence intervals for McDonald's ω were not estimated, as bootstrap-based procedures proved unstable under the present data conditions.

Contingencies of Self-Worth Scale

Contingencies of self-worth were assessed using the German version (Schütz & Sellin, 2003) of the Contingencies of Self-Worth Scale (Crocker et al., 2003). The present study included the subscales Approval from Others and Virtue, each consisting of five items. Items were rated on a 5-point Likert-typed scale ranging from 1 (*strongly disagree*) to 5 (*strongly agree*). Three items were reverse-coded prior to analysis. Subscale scores were computed as the mean of the respective items, with higher scores indicating stronger contingency of self-worth in the respective domain. Example items include "I couldn't respect myself if I didn't live up to a moral code." (German original: „Ich könnte mich nicht respektieren, würde mein Verhalten moralische Prinzipien verletzen") for Virtue and "My self-esteem depends on the opinions others hold of me. (German original: „Mein Selbstwert hängt davon ab, was andere

über mich denken.”) for Approval from Others. Reliability coefficients were as follows: CSW Approval ($\omega = .86$, $\alpha = .85$), CSW Virtue ($\omega = .76$, $\alpha = .76$).

HEXACO Personality Inventory-Revised (HEXACO-PI-R)

Furthermore, personality traits were assessed using the German version of the HEXACO-PI-R (Lee & Ashton, 2018). The present study included six facet scales: Gentleness and Flexibility (Agreeableness domain), Modesty and Sincerity (Honesty-Humility domain), and Perfectionism and Prudence (Conscientiousness domain). Each facet consisted of four items. Items were rated on a 5-point Likert-typed scale ranging from 1 (*strongly disagree*) to 5 (*strongly agree*). Twelve items were reverse-coded prior to analysis according to the scoring guidelines. Facet scores were calculated as the mean of the corresponding items. Example items include “I prefer to do whatever comes to mind, rather than stick to a plan.” (German original: „Ich ziehe es vor, das zu tun, was mir gerade in den Sinn kommt, anstatt an einem Plan festzuhalten.”; Prudence; reverse-coded) and “I wouldn't use flattery to get a raise or promotion at work, even if I thought it would succeed.” (German original: „Ich würde keine Schmeicheleien benutzen, um eine Gehaltserhöhung zu bekommen oder befördert zu werden, auch wenn ich wüsste, dass es erfolgreich wäre.”; Sincerity; Lee & Ashton, 2018). Internal consistency estimates varied from poor to acceptable: for Modesty $\omega = .66$, $\alpha = .63$, for Sincerity $\omega = .65$, $\alpha = .61$, for Gentleness $\omega = .68$, $\alpha = .65$, for Flexibility $\omega = .50$, $\alpha = .49$, for Perfectionism $\omega = .81$, $\alpha = .79$, for Prudence $\omega = .67$, $\alpha = .66$.

International Cognitive Ability Resource (ICAR)

General cognitive ability was assessed using a short form of the International Cognitive Ability Resource (ICAR; Condon & Revelle, 2014). The short form comprised four types of cognitive tasks: matrix reasoning, verbal reasoning, three-dimensional rotation, and letter-number series, with four items per task type. All items were presented in multiple-choice format. Most tasks included six response options, whereas three-dimensional rotation items included eight response options. Responses were scored dichotomously (0 = incorrect, 1 = correct), and a total cognitive ability score was computed as the sum of correct responses. Reliability analysis indicated adequate internal coherence ($\alpha = .78$).

Original OA-TB 75 Subtests

T8/1: Criticalness of Evaluation. This subtest consists of 20 items in which participants evaluate various performances. Responses are given on a 5-point scale ranging

from 1 (*very weak*) to 5 (*very good*). Scores are summed across items, with higher scores indicating lower criticalness in evaluating others' performance. An example item is: "How do you evaluate the performance of a person who writes three one-sided letters in one hour?" (German original: „Wie beurteilen Sie die Leistung einer Person, die in einer Stunde drei einseitige Briefe schreibt?"; Häcker et al., 1975). Internal consistency was good $\omega = .85$, $\alpha = .84$.

T18/1: Aphorism Acceptance. This subtest comprises 25 items in which participants indicate whether the use of a given aphorism would leave a positive impression. Response options are dichotomous („meaningless, doesn't make a good impression" vs. „meaningful, makes a good impression"). Scores reflect the total number of endorsements of the positive impression option, with higher scores indicating greater acceptance of aphoristic statements. An example item is: "As you call into the forest, so it echoes back" (as in "You reap what you sow; German original: „Wie man in den Wald hineinruft, so schallt es heraus."; Häcker et al., 1975). Examination of internal coherence revealed questionable values: $\omega = .64$, $\alpha = .57$.

T111/1: Familiar vs. Strange Story Preferences. Participants are presented with eight groups of four short story beginnings and are asked to rank them from 1 (*most interested in hearing the rest of the story*) to 4 (*least interested in hearing the rest of the story*). Each group contains two unfamiliar or unusual stories. One sample item is: „Many supernatural phenomena cannot simply be dismissed. An example for that is..." (German original: „Viele übersinnliche Erscheinungen kann man nicht einfach ablegen. Ein Beispiel dafür ist ... "; Häcker et al., 1975). Scores are computed based on the ranking of the unfamiliar stories, with higher scores indicating stronger preference for unusual content. Minor wording adaptations were made to modernize outdated terminology while preserving content. Internal coherence of the selected OA-TB subtest was questionable: $\omega = .66$, $\alpha = .57$.

T9/1: Opinion Inventory. This subtest consists of 37 items assessing agreement with various statements. Five items were excluded due to content related to religious beliefs or life after death. Minor wording adjustments were made to remove outdated or gender-specific phrasing while preserving semantic meaning. Responses are given on a 5-point agreement scale. Scoring assigns 1 point to option 4 "somewhat agree" and 2 points to option 5 "strongly agree," with higher total scores indicating stronger agreement tendencies. An

example item is “Every person is capable of performing very well in certain areas” (German original: „Jeder Mensch ist auf bestimmten Gebieten zu sehr guten Leistungen fähig.“; Häcker et al., 1975). Internal consistency was acceptable: $\omega = .74$, $\alpha = .75$.

Analysis Plan

Prior to data collection the present study was preregistered on the Open Science Framework (available at <https://osf.io/nsyqg/>). All statistical analyses were conducted using R version 4.5.2 ([Not] Part in a Rumble; R Core Team, 2025) within RStudio version 2026.01.1+403 (Apple Blossom; Posit Team, 2026). Unless otherwise specified, tests were two-tailed with an alpha level of .05, and effect sizes with 95% confidence intervals are reported where appropriate.

Prior to hypothesis testing, data were screened for missing values, distributional characteristics, and suitability for factor analysis. Factorability of the subtest scores was evaluated using the Kaiser-Meyer-Olkin (KMO; Kaiser, 1970) measure and Bartlett’s test of sphericity (Bartlett, 1951). Subtest intercorrelations were inspected to determine whether shared variance justified latent modeling. Analyses of internal structure were conducted at the subtest level, whereas correlational validity analyses were conducted at the composite score level.

To examine the underlying structure of the six newly developed subtests, an exploratory factor analysis (EFA) was performed with the package psych version 2.5.6 (Revelle, 2025) using principal axis factoring with oblimin rotation, allowing for correlated factors. The number of factors to retain was determined based on parallel analysis and inspection of the scree plot. Factor loadings of .30 or higher were considered salient, and cross-loadings were evaluated with respect to interpretability.

To further evaluate the hypothesized two-factor structure, a confirmatory factor analysis (CFA) through the package lavaan version 0.6-20 (Rosseel, 2012; Rosseel et al., 2025) was estimated in which each latent factor was defined by its three corresponding subtests representing approval-based and principle-based self-worth regulation. Robust maximum likelihood estimation (MLR) was used to account for potential deviations from normality. Model fit was evaluated according to established guidelines, including a χ^2/df ratio below 3 (Kline, 2016), CFI and TLI values $\geq .90$ ($\geq .95$ indicating good fit), RMSEA values $\leq .08$ ($\leq .05$ indicating good fit), and SRMR values $\leq .08$ (Browne & Cudeck, 1993; Hu & Bentler,

1999). A one-factor alternative model was estimated for comparison. Given the available sample size, CFA results are interpreted cautiously and primarily as complementary to an essentially exploratory examination. Internal consistency was assessed for each subtest and for both composite scores using McDonald's omega (McDonald, 1999) as the primary reliability coefficient and Cronbach's alpha (Cronbach, 1951) for comparability with previous research. Values equal to or higher than .70 were considered acceptable for early-stage instrument development. Correlations were computed using pairwise complete observations.

Convergent validity was examined using Pearson correlations between the composite scores of the newly developed tests and theoretically related external measures. Discriminant validity was evaluated through correlations with general cognitive ability (ICAR; Condon & Revelle, 2014).

To examine correspondence with the original Objective-Analytic Test Battery (OA-TB 75; Häcker et al., 1975), a factor-analytic approach was applied to the four included OA-TB subtests (T8, T18, T111, T9) using principal axis factoring with oblimin rotation.

Associations between the newly developed composites and the original OA-TB subtests were examined using Pearson correlations.

All analyses not specified a priori are reported as exploratory. Any deviations from the preregistered analysis plan are documented transparently in the Results section.

Results

Structural Analyses

Of the 89 participants included in the final sample, a total of 82 provided complete subtest scores and were retained for structural analyses following application of the predefined exclusion criteria. Descriptive statistics for the six newly developed subtests and the two composite scores are presented in Table 2:

Table 2*Descriptive Statistics of the Developed Subtests and Tests*

Variable	<i>M</i>	<i>SD</i>	Median	Min	Max
Approval-Based Self-Worth Test	2.97	0.25	3.00	2.38	3.48
SF-Subtest	2.97	0.41	3.00	2.00	3.86
CA-Subtest	3.36	0.48	3.43	2.14	4.00
TP-Subtest	2.59	0.46	2.57	1.29	3.86
Principle-Based Self-Worth Test	2.85	0.28	2.86	2.05	3.67
SR-Subtest	3.06	0.53	3.14	1.71	4.00
MS-Subtest	2.57	0.44	2.43	1.29	3.57
PR-Subtest	2.92	0.36	3.00	1.71	3.57

Note. $N = 82$; SF = Feedback Sensitivity Subtest, CA = Conformity Appraisal Subtest, and TP = Task Preference Subtests belong to Approval-Based Self-Worth Regulation Test, while SR = Structured Self-Control, MS = Mistake Sensitivity Subtest, and PR = Principles Subtests belong to the Principle-Based Self-Worth Regulation Test

Across subtests, mean scores ranged from $M = 2.57$ (Mistake Sensitivity) to $M = 3.36$ (Conformity Appraisal) on the 4-point response scale. Standard deviations ranged from $SD = 0.36$ to $SD = 0.53$, suggesting variability in responses, with no clear indication of floor or ceiling effects. Composite scores showed means of $M = 2.97$ ($SD = 0.25$) for approval-based self-worth regulation and $M = 2.85$ ($SD = 0.28$) for principle-based self-worth regulation. Intercorrelations among the six subtests were small and heterogeneous (see Table 3), ranging from $r = -.41$ to $r = .34$, with several associations close to zero.

Table 3*Intercorrelations (Pearson's r) Among Developed Subtests*

Variable	1	2	3	4	5	6
1. SF-Subtest	—	0.15	−0.15	−0.18	−0.19	−0.16
2. CA-Subtest		—	−0.09	−0.25	0.11	−0.41
3. TP-Subtest			—	0.12	0.24	0.04
4. SR-Subtest				—	−0.05	0.34
5. MS-Subtest					—	0.00
6. PR-Subtest						—

Note. $N = 82$ SF = Feedback Sensitivity Subtest, CA = Conformity Appraisal Subtest, and TP = Task Preference Subtests belong to Approval-Based Self-Worth Regulation Test, while SR = Structured Self-Control, MS = Mistake Sensitivity Subtest, and PR = Principles Subtests belong to the Principle-Based Self-Worth Regulation Test

The correlation between the two composite scores was $r = -.20$, indicating a small negative association between approval-based and principle-based regulation tendencies. Prior to exploratory factor analysis, factorability of the six subtest scores was examined. The Kaiser-Meyer-Olkin measure indicated marginal sampling adequacy ($KMO = .60$; Kaiser, 1970), with individual MSA values ranging from .46 to .67. Bartlett's test of sphericity (Bartlett, 1951) was significant, $\chi^2_{(15)} = 41.44, p < .001$, indicating that the correlation matrix differed from an identity matrix and was therefore suitable for factor analysis.

An exploratory factor analysis (EFA) using principal axis factoring with oblimin rotation was conducted on the six subtest scores. Parallel analysis suggested a predominantly unidimensional structure, as the second empirical eigenvalue only marginally exceeded the corresponding simulated eigenvalue, indicating weak support for a multidimensional solution. Given the theoretical expectation of two higher-order regulation styles, both one-factor and two-factor solutions were examined. The two-factor solution accounted for 31.1% of the total variance (Factor 1: 19.1%; Factor 2: 12.0%). However, the loading pattern did not clearly reflect the theoretically proposed distinction between approval-based and principle-based regulation. Factor 1 showed substantial loadings for the subtests Principles (PR; .66) and Structured Reasoning (SR; .50), and a negative loading for Conformity Appraisal (CA; −

.59), with Feedback Sensitivity (SF; $-.30$) loading modestly. Factor 2 was primarily defined by Mistake Sensitivity (MS; $.71$) and Task Preference (TP; $.35$), with SF additionally showing a secondary loading ($-.29$). Several subtests exhibited cross-loadings or loadings inconsistent with their theoretical assignment. The correlation between the two extracted factors was negligible ($r = -.04$).

The one-factor solution accounted for 18.8% of the total variance. In this solution, PR ($.66$) and SR ($.50$) loaded positively, whereas CA ($-.57$) and SF ($-.30$) loaded negatively. TP and MS showed minimal loadings. Overall model fit for the one-factor solution was acceptable (RMSEA = $.06$), although the proportion of explained variance remained limited.

Taken together, the EFA results did not provide clear empirical support for the hypothesized two-factor structure. While a two-factor solution could be statistically extracted, the loading pattern was weak and only partially aligned with theoretical expectations. These findings are consistent with the modest intercorrelations observed among the subtests. Based on these results, confirmatory factor analyses were conducted to formally compare a unidimensional model with the theoretically specified two-factor model.

A confirmatory factor analysis (CFA) was conducted to test the preregistered two-factor model in which approval-based self-worth regulation was defined by Social Feedback (SF), Conformity Appraisal (CA), and Task Preference (TP), and principle-based self-worth regulation was defined by Self-Regulation (SR), Mistake Sensitivity (MS), and Principled Reasoning (PR). A unidimensional alternative model was estimated for comparison.

The two-factor model yielded $\chi^2_{(8)} = 12.08$, $p = .15$, with CFI = $.86$ (robust CFI = $.84$), TLI = $.73$ (robust TLI = $.71$), RMSEA = $.08$ (robust RMSEA = $.08$, 90% CI $[.00, .16]$), and SRMR = $.08$. Standardized factor loadings were heterogeneous. Within the approval-based factor, CA loaded moderately ($\lambda = .61$), whereas SF ($\lambda = .28$) and TP ($\lambda = -.15$) showed weak loadings. Within the principle-based factor, PR ($\lambda = .69$) and SR ($\lambda = .49$) loaded moderately, whereas MS exhibited a negligible loading ($\lambda = -.03$). The estimated latent correlation between the two factors was very high ($r = -.94$).

The one-factor model produced comparable fit indices, $\chi^2_{(9)} = 12.12$, $p = .21$, CFI = $.89$ (robust CFI = $.88$), TLI = $.82$ (robust TLI = $.81$), RMSEA = $.07$ (robust RMSEA = $.07$, 90% CI $[.00, .15]$), and SRMR = $.08$. Standardized loadings ranged from negligible (MS: $\lambda = .02$) to moderate (PR: $\lambda = -.68$; CA: $\lambda = .58$; SR: $\lambda = -.49$), whereas SF ($\lambda = .28$) and TP ($\lambda =$

-.14) showed weak loadings. Fit indices were similar across models, with the unidimensional solution showing slightly higher CFI and TLI values. Given the extremely high latent correlation in the two-factor model and the absence of clearly differentiated loading patterns, the data did not provide strong empirical support for two distinct higher-order regulation factors in this sample.

Convergent and Discriminant Validity

Pearson correlations were computed using pairwise complete observations to examine convergent and discriminant validity. Sample sizes varied from $N = 69$ to $N = 82$ across analyses due to differential missingness in the external measures. As hypothesized, the Approval-Based Self-Worth regulation Test showed a positive association with the CSW Approval from Others (Crocker et al., 2003) subscale, $r = .37$, 95% CI [.16, .55], $p < .001$, indicating moderate convergent validity with externally contingent self-worth. In contrast, the Principle-Based Self-Worth Regulation Test was not significantly related to the CSW Virtue (Crocker et al., 2003) subscale, $r = .11$, 95% CI [-.11, .32], $p = .33$. Thus, convergent validity was supported for approval-based regulation but not clearly supported for principle-based regulation with respect to the CSW measure.

Regarding associations with HEXACO facets, the Approval-Based Self-Worth regulation test was not significantly related to Gentleness ($r = .16$, 95% CI [-.06, .36], $p = .16$), Flexibility ($r = -.02$, 95% CI [-.24, .20], $p = .83$), or Modesty ($r = -.05$, 95% CI [-.27, .18], $p = .68$). The Principle-Based Self-Worth Regulation Test showed positive associations with Perfectionism ($r = .28$, 95% CI [.07, .47], $p = .01$) and Prudence ($r = .26$, 95% CI [.05, .45], $p = .02$), whereas its association with Sincerity did not reach statistical significance ($r = -.19$, 95% CI [-.39, .03], $p = .08$). Overall, these findings provide partial support for theoretically expected associations between principle-based regulation and conscientiousness-related HEXACO facets.

With respect to cognitive ability, the Approval-Based Self-Worth Regulation Test was not significantly associated with ICAR (Condon & Revelle, 2014) total performance ($r = -.10$, 95% CI [-.33, .14], $p = .43$). The Principle-Based Self-Worth Regulation Test showed a negative association with ICAR performance ($r = -.25$, 95% CI [-.46, -.02], $p = .04$). Although statistically significant, the effect size was small to medium, suggesting limited overlap between principle-based regulation and general cognitive ability.

Finally, associations with the selected original OA-TB (Häcker et al., 1975) subtests were examined. The Approval-Based Self-Worth Regulation Test was not significantly related to T8 (BL; $r = .13$, 95% CI $[-.11, .35]$, $p = .29$), T18 (RW; $r = -.02$, 95% CI $[-.26, .22]$, $p = .85$), or T111 (GE; $r = -.16$, 95% CI $[-.38, .08]$, $p = .20$). The Principle-Based Self-Worth Regulation Test was not significantly associated with T9 (ME; $r = -.02$, 95% CI $[-.25, .22]$, $p = .88$). Overall, associations between the developed subtests and the selected OA-TB subtests were limited in the present sample.

Original OA-TB Subtests Structure

To examine structural relations among the included OA-TB 75 subtests (T8/BL, T18/RW, T111/GE, and T9/ME; Häcker et al., 1975), exploratory factor analyses were conducted. Preliminary diagnostics indicated limited factorability. The Kaiser-Meyer-Olkin measure of sampling adequacy was low (KMO = .42; Kaiser, 1970), and Bartlett's test of sphericity (Bartlett, 1951) was not significant, $\chi^2_{(6)} = 10.04$, $p = .12$, suggesting insufficient shared variance among the subtests to justify a stable factor solution. Given the limited factorability indices, the factor analyses are interpreted as exploratory and descriptive in nature. A one-factor solution accounted for 15.0% of the total variance. RW ($\lambda = .54$) and ME ($\lambda = .55$) showed moderate loadings on this factor, whereas BL and GE did not load substantially. Model fit was weak (RMSEA = .11), indicating limited support for a unidimensional structure. A two-factor solution was also explored; however, estimation did not converge (maximum number of 50 iterations exceeded), indicating model instability. Although RW and ME loaded strongly on one factor, the remaining subtests did not form a coherent second factor. Given the low factorability indices and convergence problems, the two-factor solution was not considered interpretable.

Discussion

The present study aimed to develop and provide an initial validation of objective subtests designed to operationalize two source traits described by Cattell and Warburton (1967): U.I. 29 and U.I. 36. Based on the original behavioral descriptions, these traits were conceptually translated into contemporary terms as approval-based self-worth regulation (U.I. 29) and principle-based self-worth regulation (U.I. 36). While conceptually related, the two constructs were theorized to reflect distinct regulatory orientations: one organized around external evaluation and social approval, and the other around internally endorsed standards

and the maintenance of self-respect. This distinction aligns with broader theoretical frameworks. The contingencies-of-self-worth model (Crocker & Wolfe, 2001) differentiates externally contingent self-worth (e.g., approval from others) from internally grounded contingencies (e.g., virtue). The present study therefore sought to examine whether these theoretically distinct orientations could be captured through objective assessment formats in the spirit of Cattell's original proposal.

The structural analyses did not provide clear empirical support for the hypothesized two-factor differentiation between approval-based and principle-based self-worth regulation. Hypotheses 1 and 2, which predicted clustering of the six subtests into two higher-order dimensions and superior fit of a two-factor model relative to a unidimensional structure, were not supported. Exploratory analyses suggested only limited evidence for a two-dimensional structure, and confirmatory analyses yielded comparable model fit for one- and two-factor solutions. Most notably, the extremely high negative latent correlation ($r = -.94$) between the two factors in the two-factor CFA model indicated near-linear dependency rather than clearly separable constructs, contrary to Hypothesis 3. Although this pattern may partly reflect sample size limitations and the aggregation of indicators at the subtest level, it also suggests that the present operationalization may not sufficiently differentiate between externally guided and internally guided regulatory tendencies at the behavioral level. Several subtests showed heterogeneous or weak factor loadings, and reliability estimates were modest for multiple scales. Thus, Hypothesis 4, which predicted adequate internal consistency for all subtests and composite dimensions, was only partially supported.

A closer examination of item characteristics provides further insight. In some subtests, response options may have been conceptually similar or insufficiently distinct in their motivational implications. When alternatives differ only subtly in wording or implied regulatory stance, respondents may perceive them as interchangeable, reducing internal coherence. In consistence with this interpretation, alpha-if-item-deleted values did not indicate that removing any individual item substantially improved internal consistency. In addition, the forced-choice, scenario-based format employed in the present study may have been relatively unfamiliar to participants. Unlike traditional Likert-type self-report items, situationally embedded response formats require interpretation of context, inference of implicit motives, and selection among behaviorally nuanced alternatives. This format shares

features with situational judgment tests (SJTs), which present respondents with realistic scenarios and ask them to evaluate or select behavioral responses (Lievens et al., 2008; Whetzel & McDaniel, 2009). Although SJTs have demonstrated incremental validity and robustness in applied contexts, their psychometric properties can vary substantially depending on response format, scoring method, and clarity of response alternatives (Kepes et al., 2025; Robb et al., 2026). Recent research emphasizes that subtle differences between response options and ambiguous behavioral may contribute to modest internal consistency and unclear factor structures (Whetzel et al., 2020). The present findings may reflect similar challenges inherent in developing behavior-based assessment formats that attempt to infer stable regulatory tendencies from scenario judgments. Moreover, because the present subtests were designed to obscure the construct being measured—consistent with the definition of objective testing (Cattell & Warburton, 1967)—items were intentionally indirect and avoided explicit trait language. While this approach strengthens resistance to self-presentation biases, it may also reduce internal transparency and, consequently, inter-item homogeneity in early development stages.

Taken together, the structural findings suggest that the development of the present objective personality test versions require further refinement of item wording, response formats, and behavioral operationalization. The absence of clear two-factor differentiation in this initial validation does not necessarily contradict the theoretical distinction between externally and internally grounded regulation. Rather, it indicates that the behavioral translation of these constructs remains psychometrically underdeveloped and warrants further calibration.

Associations with Self-Report Measures and General Cognitive Ability

Despite limited structural differentiation at the subtest level, associations with external measures were examined to evaluate the interpretability of the composite scores representing approval-based and principle-based self-worth orientations. Consistent with Hypothesis 5, the approval-based composite showed a moderate positive association with the CSW Approval from Others (Crocker et al., 2003) subscale. This pattern supports the intended interpretation of this score as reflecting a trait-like orientation toward externally contingent self-worth. In contrast, Hypothesis 6 was not supported, as the principle-based composite did not show a significant association with the CSW Virtue subscale (Crocker et al., 2003). This discrepancy may indicate that the behavioral tendencies captured by the objective tasks do

not directly map onto explicit self-reported moral self-evaluations. With respect to broader personality traits, expected associations were only weakly supported. The approval-based orientation did not demonstrate the predicted relations with the HEXACO facets Gentleness, Flexibility, or Modesty (Lee & Ashton, 2018). In contrast, the principle-based orientation showed positive associations with the facets Perfectionism and Prudence, suggesting convergence with conscientiousness-related tendencies characterized by disciplined and controlled behavior. As anticipated, associations with cognitive ability were small in magnitude and partially non-significant, supporting discriminant validity from intellectual performance.

Original OA-TB Subtests

The expected historical correspondence between the newly developed composites and selected OA-TB indicators was not clearly observed. The included OA-TB subtests (Häcker et al., 1975) did not cluster into a stable structure, and exploratory analyses revealed low factorability and convergence difficulties. This pattern must be interpreted in light of the original development context of Cattell's objective factors (Cattell & Warburton, 1967). The source traits reported in the compendium were derived from extensive factor-analytic investigations across large pools of behavioral variables and numerous objective tests (Cattell, 1958; Cattell & Warburton, 1967). Importantly, factor extraction was based on hundreds of variables rather than a small set of aggregated test scores. It was explicitly emphasized that broad variable sampling is essential for stable factor identification within the objective data domain (Cattell & Warburton, 1967). In the present study, only four OA-TB subtests (Häcker et al., 1975) were included, and analyses were conducted at the subtest level rather than at the behavioral-variable level. Under these conditions, the recovery of a stable latent structure is statistically unlikely (MacCallum et al., 1999). Factor analytic results derived from such a restricted indicator set cannot be considered equivalent to the large-scale variable matrices on which the original source traits were based.

Implications

The present findings have implications for renewed efforts to develop objective personality tests within Cattell and Warburton's (1967) original framework. Historically, OPTs were intended as cumulative, behavior-based indicators of stable source traits derived from extensive variable sampling and iterative factor analysis (Cattell, 1958; Cattell & Warburton,

1967). However, subsequent developments in personality assessment have largely shifted toward self-report methodologies, while objective approaches have either been applied to isolated constructs or developed independently of Cattell's taxonomy (Romero et al., 2024). As a result, the original OPT tradition has not been systematically extended. Modern behavior-based personality measures are rarely constructed as direct conceptual continuations of Cattell's objective source traits. One possible reason for the limited uptake of OPTs is the difficulty of establishing strong internal structure and validity evidence within small-scale studies. The present findings illustrate these challenges: structural differentiation was limited, and reliability varied across subtests. However, this does not necessarily imply that the underlying constructs lack coherence. Rather, it underscores the methodological demands of objective assessment, which requires broad behavioral sampling and cumulative validation rather than isolated test construction. Importantly, the original objective factors (Cattell & Warburton, 1967) were derived from large behavioral matrices, not from small sets of aggregated test scores. The modest structural clarity observed here may therefore reflect restricted indicator breadth rather than conceptual inadequacy. If objective personality testing has remained uncommon, the reason may lie less in theoretical weakness and more in the resource-intensive nature of its proper development (Ortner & Proyer, 2015; Romero et al., 2024). The partial convergent validity findings provided limited support that behavior-based operationalizations of self-worth-related personality tendencies are feasible. Future work should expand behavioral indicators, increase sample size, and pursue systematic, cumulative refinement through iterative test development and validation (Clark & Watson, 1995).

Limitations

Several limitations should be considered when interpreting the present findings. First, the sample size was modest for structural modeling, which may have reduced the stability and precision of parameter estimates. Structural analyses were conducted at the aggregated subtest level. Item-level factor analysis would have required a substantially larger sample to obtain stable parameter estimates. Given the present sample size, aggregation represented a statistically appropriate compromise for initial structural evaluation. Nevertheless, this approach may have limited the ability to detect more fine-grained latent differentiation among behavioral indicators. In addition, internal consistency varied across subtests,

suggesting that certain scenarios and response alternatives require further refinement to achieve clearer construct differentiation.

Finally, the cross-sectional design does not permit conclusions regarding temporal stability. Replication in larger samples, expanded behavioral sampling, and item-level structural analyses will be necessary to establish more robust psychometric evidence and clarify the dimensional structure of the proposed constructs.

Conclusion

The present study sought to re-examine and empirically evaluate Objective Personality Tests in the tradition of Cattell (1958) by developing and initially validating behavior-based subtests for U.I. 29 and U.I. 36. Drawing on Cattell and Warburton's (1967) original conceptualization and integrating contemporary perspectives on self-worth regulation, the study translated these source traits into structured, scenario-based behavioral formats suitable for standardized assessment.

The hypothesized two-factor structure was not clearly supported, and structural differentiation between approval-based and principle-based orientations remained limited. Furthermore, convergent validity findings offer only ambiguous evidence that the developed subtests capture personality-related tendencies. Further refinement and validation across multiple studies would be required before firm conclusions can be drawn, and even such efforts would not necessarily guarantee recovery of the hypothesized structures. More broadly, the findings illustrate both the general demands of psychological test development and the stringent methodological requirements of objective personality assessment. The continued dominance of self-report measures in personality assessment is supported by their robust psychometric properties and predictive validity. Nevertheless, self-report instruments primarily capture individuals' self-perceptions and identity-based narratives. Revisiting the objective personality testing approach (Cattell & Warburton, 1967) offers a theoretically grounded pathway for expanding personality assessment beyond introspective self-description. Objective behavioral indicators may provide complementary information by assessing how trait-relevant tendencies are enacted in structured situations, thereby extending rather than replacing questionnaire-based approaches.

References

- Anglim, J., Horwood, S., Smillie, L. D., Marrero, R. J., & Wood, J. K. (2020). Predicting psychological and subjective well-being from personality: A meta-analysis. *Psychological Bulletin*, 146(4), 279–323. <https://doi.org/10.1037/bul0000226>
- Ashton, M. C., & Lee, K. (2008). The HEXACO Model of personality structure. In G. Boyle, G. Matthews, & D. Saklofske, *The SAGE Handbook of Personality Theory and Assessment: Volume 2—Personality Measurement and Testing* (pp. 239–260). SAGE Publications Ltd. <https://doi.org/10.4135/9781849200479.n12>
- Ashton, M. C., & Lee, K. (2017). HEXACO Personality Inventory-Revised (HEXACO-PI-R). In V. Zeigler-Hill & T. K. Shackelford (Eds.), *Encyclopedia of Personality and Individual Differences* (pp. 1–3). Springer International Publishing. https://doi.org/10.1007/978-3-319-28099-8_900-1
- Ashton, M. C., & Lee, K. (2025). Self- and observer reports of personality. *Annual Review of Psychology*, 76(1), 771–795. <https://doi.org/10.1146/annurev-psych-020124-115044>
- Bartlett, M. S. (1951). The Effect of standardization on a χ^2 approximation in factor analysis. *Biometrika*, 38(3/4), 337. <https://doi.org/10.2307/2332580>
- Browne, M. W., & Cudeck, R. (1993). Alternative ways of assessing model fit. In K. A. Bollen & J. S. Long (Eds.), *Testing structural equation models* (pp. 136–162). New-bury Park, CA: Sage.
- Cattell, R. B. (1948). Primary personality factors in the realm of objective tests. *Journal of Personality*, 16(4), 459–486. <https://doi.org/10.1111/j.1467-6494.1948.tb02301.x>
- Cattell, R. B. (1955). The principal replicated factors discovered in objective personality tests. *The Journal of Abnormal and Social Psychology*, 50(3), 291–314. <https://doi.org/10.1037/h0046312>

- Cattell, R. B. (1958). What is “objective” in “objective personality tests?” *Journal of Counseling Psychology*, 5(4), 285–289. <https://doi.org/10.1037/h0046268>
- Cattell, R. B., & Warburton, F. W. (1967). *Objective personality and motivation tests: A theoretical introduction and practical compendium*. University of Illinois Press.
- Chen, S., Cheung, A. C. K., & Zeng, Z. (2025). Big Five personality traits and university students’ academic performance: A meta-analysis. *Personality and Individual Differences*, 240, 113163. <https://doi.org/10.1016/j.paid.2025.113163>
- Clark, L. A., & Watson, D. (1995). Constructing validity: Basic issues in objective scale development. *Psychological Assessment*, 7(3), 309–319. <https://doi.org/10.1037/1040-3590.7.3.309>
- Corneille, O., & Gawronski, B. (2024). Self-reports are better measurement instruments than implicit measures. *Nature Reviews Psychology*, 3(12), 835–846. <https://doi.org/10.1038/s44159-024-00376-z>
- Condon, D. M., & Revelle, W. (2014). The international cognitive ability resource: Development and initial validation of a public-domain measure. *Intelligence*, 43, 52–64. <https://doi.org/10.1016/j.intell.2014.01.004>
- Crocker, J., Brook, A. T., Niiya, Y., & Villacorta, M. (2006). The pursuit of self-esteem: Contingencies of self-worth and self-regulation. *Journal of Personality*, 74(6), 1749–1771. <https://doi.org/10.1111/j.1467-6494.2006.00427.x>
- Crocker, J., Luhtanen, R. K., Cooper, M. L., & Bouvrette, A. (2003). Contingencies of self-worth in college students: Theory and measurement. *Journal of Personality and Social Psychology*, 85(5), 894–908. <https://doi.org/10.1037/0022-3514.85.5.894>

- Crocker, J., & Wolfe, C. T. (2001). Contingencies of self-worth. *Psychological Review*, 108(3), 593–623. <https://doi.org/10.1037/0033-295X.108.3.593>
- Cronbach, L. J. (1951). Coefficient alpha and the internal structure of tests. *Psychometrika*, 16:297-334.
- Deci, E. L., & Ryan, R. M. (1985). *Intrinsic motivation and self-determination in human behavior*. Plenum press.
- Donnellan, M. B., Trzesniewski, K. H., & Robins, R. W. (2015). Measures of self-esteem. *Measures of Personality and Social Psychological Constructs* (pp. 131–157). Elsevier. <https://doi.org/10.1016/B978-0-12-386915-9.00006-1>
- Faul, F., Erdfelder, E., Lang, A.-G., & Buchner, A. (2007). G*Power 3: A flexible statistical power analysis program for the social, behavioral, and biomedical sciences. *Behavior Research Methods*, 39, 175-191.
- Faul, F., Erdfelder, E., Buchner, A., & Lang, A.-G. (2009). Statistical power analyses using G*Power 3.1: Tests for correlation and regression analyses. *Behavior Research Methods*, 41, 1149-1160
- Greenwald, A. G., & Farnham, S. D. (2000). Using the Implicit Association Test to measure self-esteem and self-concept. *Journal of Personality and Social Psychology*, 79(6), 1022–1038. <https://doi.org/10.1037/0022-3514.79.6.1022>
- Greenwald, A. G., McGhee, D. E., & Schwartz, J. L. K. (1998). Measuring individual differences in implicit cognition: The implicit association test. *Journal of Personality and Social Psychology*, 74(6), 1464–1480. <https://doi.org/10.1037/0022-3514.74.6.1464>
- Häcker, H., Schmidt, L. R., Schwenkmezger, P., & Utz, H. E. (1975). *Objektive Testbatterie OA-TB 75 [Testmanual]*. Beltz Test.

- Hergovich, A., Arendasy, M. E., Sommer, M., & Bognar, B. (2007). The Vienna Risk-Taking Test--Traffic: A new measure of road traffic risk-taking. *Journal of Individual Differences*, 28(4), 198–204. <https://doi.org/10.1027/1614-0001.28.4.198>
- Hoerger, M. (2010). Participant dropout as a function of survey length in internet-mediated university studies: Implications for study design and voluntary participation in psychological research. *Cyberpsychology, Behavior, and Social Networking*, 13(6), 697–700. <https://doi.org/10.1089/cyber.2009.0445>
- Howard, J. L., Bureau, J., Guay, F., Chong, J. X. Y., & Ryan, R. M. (2021). Student motivation and associated outcomes: A meta-analysis from self-determination theory. *Perspectives on Psychological Science*, 16(6), 1300–1323. <https://doi.org/10.1177/1745691620966789>
- Hu, L., & Bentler, P. M. (1999). Cutoff criteria for fit indexes in covariance structure analysis: Conventional criteria versus new alternatives. *Structural Equation Modeling*, 6(1), 1–55. <https://doi.org/10.1080/10705519909540118>
- Johnson, J. A. (1997). Units of analysis for the description and explanation of personality. In *Handbook of Personality Psychology* (pp. 73–93). Elsevier. <https://doi.org/10.1016/B978-012134645-4/50004-4>
- Joshanloo, M. (2022). Longitudinal relationships between personality traits and social well-being: A two-decade study. *Journal of Happiness Studies*, 23(6), 2969–2983. <https://doi.org/10.1007/s10902-022-00534-1>
- Jusepeitis, A., & Rothermund, K. (2022). No elephant in the room: The incremental validity of implicit self-esteem measures. *Journal of Personality*, 90(6), 916–936. <https://doi.org/10.1111/jopy.12705>

- Kaiser, H. F. (1970). A Second Generation Little Jiffy. *Psychometrika*, 35(4), 401–415.
<https://doi.org/10.1007/BF02291817>
- Kepes, S., Keener, S. K., Lievens, F., & McDaniel, M. A. (2025). An integrative, systematic review of the situational judgment test literature. *Journal of Management*, 51(6), 2278–2319.
<https://doi.org/10.1177/01492063241288545>
- Kim, H., Di Domenico, S. I., & Connelly, B. S. (2019). Self–other agreement in personality reports: A Meta-analytic comparison of self- and informant-report means. *Psychological Science*, 30(1), 129–138. <https://doi.org/10.1177/0956797618810000>
- Kline, R. B. (2016). *Principles and practice of structural equation modeling* (Fourth edition). The Guilford Press.
- Koch, T., Ortner, T. M., Eid, M., Caspers, J., & Schmitt, M. (2014). Evaluating the construct validity of Objective Personality Tests using a Multitrait-Multimethod-Multioccasion- (MTMM-MO)-approach. *European Journal of Psychological Assessment*, 30(3), 208–230.
<https://doi.org/10.1027/1015-5759/a000212>
- Krause, S., Back, M. D., Egloff, B., & Schmukle, S. C. (2011). Reliability of implicit self–esteem measures revisited. *European Journal of Personality*, 25(3), 239–251.
<https://doi.org/10.1002/per.792>
- Kubinger, K. D. & Ebenhöf, J. (1996). *Arbeitshaltungen - Kurze Testbatterie: Anspruchsniveau, Frustrationstoleranz, Leistungsmotivation, Impulsivität/Reflexivität [Work Styles – A short test-battery: Level of aspiration, achievement motivation, frustration tolerance, impulsiveness/reflexiveness]* Test: Software and Manual. Frankfurt: Swets.
- Kubinger, K. D. & Ebenhöf, J. (2002). *Arbeitshaltungen - Kurze Testbatterie: Anspruchsniveau, Frustrationstoleranz, Leistungsmotivation, Impulsivität/Reflexivität [Work Style – A short*

test-battery: Level of aspiration, achievement motivation, frustration tolerance, impulsiveness/reflexiveness] – Version 26.00. Test: Software and Manual (writer of this manual: T. Karner & M. Sommer). Mödling: Wiener Testsystem/ Schuhfried.

Kubinger, K. D. (2009). The technique of objective personality-tests sensu R. B. Cattell nowadays: The Viennese pool of computerized tests aimed at experiment-based assessment of behavior. *Acta Psychologica Sinica*, 41(10), 1024–1036.

<https://doi.org/10.3724/SP.J.1041.2009.01024>

Kyriazos, T. A. (2018). Applied psychometrics: sample size and sample power considerations in factor analysis (EFA, CFA) and SEM in general. *Psychology*, 9(8), 2207–2230.

<https://doi.org/10.4236/psych.2018.98126>

Latkin, C. A., Edwards, C., Davey-Rothwell, M. A., & Tobin, K. E. (2017). The relationship between social desirability bias and self-reports of health, substance use, and social network factors among urban substance users in Baltimore, Maryland. *Addictive Behaviors*, 73, 133–136. <https://doi.org/10.1016/j.addbeh.2017.05.005>

Lee, K., & Ashton, M. C. (2018). Psychometric properties of the HEXACO-100. *Assessment*, 25, 543-556.

Lejuez, C. W., Read, J. P., Kahler, C. W., Richards, J. B., Ramsey, S. E., Stuart, G. L., Strong, D. R., & Brown, R. A. (2002). Evaluation of a behavioral measure of risk taking: The Balloon Analogue Risk Task (BART). *Journal of Experimental Psychology: Applied*, 8(2), 75–84.

<https://doi.org/10.1037/1076-898X.8.2.75>

Lievens, F., Peeters, H., & Schollaert, E. (2008). Situational judgment tests: A review of recent research. *Personnel Review*, 37(4), 426–441. <https://doi.org/10.1108/00483480810877598>

- MacCallum, R. C., Widaman, K. F., Zhang, S., & Hong, S. (1999). Sample size in factor analysis. *Psychological Methods*, 4(1), 84–99. <https://doi.org/10.1037/1082-989X.4.1.84>
- Mammadov, S. (2022). Big Five personality traits and academic performance: A meta-analysis. *Journal of Personality*, 90(2), 222–255. <https://doi.org/10.1111/jopy.12663>
- McDonald, R. P. (1999). *Test theory: A unified treatment*. Lawrence Erlbaum Associates Publishers.
- McDonald, J. A. (2008). Measuring personality constructs: The advantages and disadvantages of self-reports, informant reports and behavioural assessments. *Enquire*.
https://www.academia.edu/79046768/Measuring_personality_constructs_The_advantages_and_disadvantages_of_self_reports_informant_reports_and_behavioural_assessments
- Olino, T. M., & Klein, D. N. (2015). Psychometric comparison of self- and informant-reports of personality. *Assessment*, 22(6), 655–664. <https://doi.org/10.1177/1073191114567942>
- Ortner, T. M., & Proyer, R. T. (2015). Objective personality tests. In *Behavior-based assessment in psychology: Going beyond self-report in the personality, affective, motivation, and social domains* (pp. 133–149). Hogrefe. <https://doi.org/10.1027/00437-000>
- Ortner, T. M., & Schmitt, M. (2014). Advances and continuing challenges in objective personality testing. *European Journal of Psychological Assessment*, 30(3), 163–168.
<https://doi.org/10.1027/1015-5759/a000213>
- Paulhus, D. L. (1984). Two-component models of socially desirable responding. *Journal of Personality and Social Psychology*, 46(3), 598–609. <https://doi.org/10.1037/0022-3514.46.3.598>
- Paulhus, D. L., & Vazire, S. (2007). The self-report method. In *Handbook of research methods in personality psychology* (pp. 224–239). The Guilford Press.

- Paulhus, D. L. (2017). Socially desirable responding on self-reports. In V. Zeigler-Hill & T. K. Shackelford (Eds.), *Encyclopedia of Personality and Individual Differences* (pp. 1–5). Springer International Publishing. https://doi.org/10.1007/978-3-319-28099-8_1349-1
- Pavlov, G., Shi, D., Maydeu-Olivares, A., & Fairchild, A. (2021). Item desirability matching in forced-choice test construction. *Personality and Individual Differences*, 183, 111114. <https://doi.org/10.1016/j.paid.2021.111114>
- Phan, L. V., & Rauthmann, J. F. (2021). Personality computing: New frontiers in personality assessment. *Social and Personality Psychology Compass*, 15(7), e12624. <https://doi.org/10.1111/spc3.12624>
- Pinon, A. (2019). Why most psychologists should assess and report personality. *Frontiers in Psychology*, 10. <https://doi.org/10.3389/fpsyg.2019.01982>
- Posit team (2026). *RStudio: Integrated Development Environment for R*. Posit Software, PBC, Boston, MA. URL <http://www.posit.co/>
- Pozzebon, J. A., & Ashton, M. C. (2009). Personality and values as predictors of self- and peer-reported behavior. *Journal of Individual Differences*, 30(3), 122–129. <https://doi.org/10.1027/1614-0001.30.3.122>
- R Core Team (2025). *R: A language and environment for statistical computing*. R Foundation for Statistical Computing, Vienna, Austria. <https://www.R-project.org/>
- Revelle, W. (2025). *psych: Procedures for psychological, psychometric, and personality research*. Northwestern University, Evanston, Illinois. R package version 2.5.6. <https://CRAN.R-project.org/package=psych>
- Robb, C., Banks, P. W., Copeland, H. L., MacIntosh, A., Ivan, R., Moskowitz, J. B., Reiter, H., & Sitarenios, G. (2026). examining the predictive validity of an open-response situational

- judgment test with typed-response and video-response items. *Educational Assessment*, 31(1), 37–47. <https://doi.org/10.1080/10627197.2025.2576228>
- Roberts, B. W., Kuncel, N. R., Shiner, R., Caspi, A., & Goldberg, L. R. (2007). The power of personality: The comparative validity of personality traits, socioeconomic status, and cognitive ability for predicting important life outcomes. *Perspectives on Psychological Science*, 2(4), 313–345. <https://doi.org/10.1111/j.1745-6916.2007.00047.x>
- Romero, M., Hernández López, J. M., & Santacreu, J. (2024). How to design an objective test to assess personality: Step by step. *Revista CES Psicología*, 17(3), 129–158. <https://doi.org/10.21615/cesp.7279>
- Rosenberg, M. (1965). *Rosenberg Self-Esteem Scale (RSES)* [Database record]. APA PsycTests. <https://doi.org/10.1037/t01038-000>
- Rosseel Y, Jorgensen T, De Wilde L (2025). *lavaan: Latent Variable Analysis*. R package version 0.6-20, <https://CRAN.R-project.org/package=lavaan>
- Rosseel, Y. (2012). lavaan: An R package for structural equation modeling. *Journal of Statistical Software*, 48(2). <https://doi.org/10.18637/jss.v048.i02>
- Ryan, R. M., & Deci, E. L. (2020). Intrinsic and extrinsic motivation from a self-determination theory perspective: Definitions, theory, practices, and future directions. *Contemporary Educational Psychology*, 61, 101860. <https://doi.org/10.1016/j.cedpsych.2020.101860>
- Santacreu, J., Rubio, V. J., & Hernández, J. M. (2006). The objective assessment of personality: Cattells's T-data revisited. *Psychology Science*, 48(1), 53–68.
- Schimmack, U., & Diener, E. (2003). Predictive validity of explicit and implicit self-esteem for subjective well-being. *Journal of Research in Personality*, 37(2), 100–106. [https://doi.org/10.1016/S0092-6566\(02\)00532-9](https://doi.org/10.1016/S0092-6566(02)00532-9)

- Schütz, A., & Sellin, I. (2003). *Skala zur Erfassung von Selbstwertkontingenzen: Deutsche Version der Contingencies of Self-Worth Scale (CSWS) von Crocker et al.* [Unpublished manuscript]. Technische Universität Chemnitz.
- Strickhouser, J. E., Zell, E., & Krizan, Z. (2017). Does personality predict health and well-being? A metasynthesis. *Health Psychology: Official Journal of the Division of Health Psychology, American Psychological Association*, 36(8), 797–810. <https://doi.org/10.1037/hea0000475>
- Weston, S. J., Hill, P. L., & Jackson, J. J. (2015). Personality traits predict the onset of disease. *Social Psychological and Personality Science*, 6(3), 309–317. <https://doi.org/10.1177/1948550614553248>
- Whetzel, D. L., & McDaniel, M. A. (2009). Situational judgment tests: An overview of current research. *Human Resource Management Review, Employee Selection at the Beginning of the 21st Century*, 19(3), 188–202. <https://doi.org/10.1016/j.hrmr.2009.03.007>
- Whetzel, D., Sullivan, T., & McCloy, R. (2020). Situational Judgment Tests: An overview of development practices and psychometric characteristics. *Personnel Assessment and Decisions*, 6(1). <https://doi.org/10.25035/pad.2020.01.001>

List of Tables

Table 1. Internal Consistency Coefficients for the Developed Subtests	23
Table 2. Descriptive Statistics of the Developed Subtests and Tests	28
Table 3. Intercorrelations (Pearson's r) Among Developed Subtests	29

List of Abbreviations

CA: Conformity Appraisal
 CSW: Contingencies of Self-Worth
 HEXACO: HEXACO Personality Inventory
 IAT: Implicit Association Test
 ICAR: International Cognitive Ability Resource
 MS: Mistake Sensitivity
 OA-TB: Objective-Analytic Test Battery
 OPT(s): Objective Personality Test(s)
 PR: Principles
 SF: Feedback Sensitivity
 SR: Structured Mental Self-Control
 TP: Task Preference
 U.I.: Universal Index

Appendix

Abstract (English)

The present study aimed to re-examine Objective Personality Tests (OPTs) in the tradition of Cattell by developing and providing an initial validation of behavior-based subtests for two objective source traits, U.I. 29 and U.I. 36. Drawing on Cattell's conceptualization of objective assessment and contemporary theories of self-worth orientation, six structured, scenario-based subtests were constructed to assess approval-based and principle-based self-worth regulation tendencies. A sample of 89 participants completed the measures, of whom 82 provided complete subtest data for structural analyses. Internal structure was examined using exploratory and confirmatory factor analyses at the subtest level. The hypothesized two-factor structure was not supported, and structural differentiation between the proposed dimensions was limited. Internal consistency was predominantly low across subtests. Convergent validity analyses showed that the approval-based self-worth regulation test was positively associated with an established measure of self-worth contingent on others' approval, while the principle-based test did not show the expected association with a measure of principle-based self-worth. Associations with additional conceptually related personality traits yielded ambiguous results, whereas divergent validity was indicated by small, partially non-significant correlations with a general cognitive ability instrument. Overall, the findings reflect the well-documented psychometric challenges of objective personality test development and highlight that cumulative refinement across multiple studies is necessary to further evaluate behavior-based indicators of self-worth regulation.

Abstract (German)

Die vorliegende Studie hatte zum Ziel, Objektive Persönlichkeitstests (OPTs) in der Tradition Cattells wiederaufzugreifen, indem verhaltensbasierte Subtests für zwei objektive Faktoren, U.I. 29 und U.I. 36, entwickelt und einer ersten Validierung unterzogen wurden. Aufbauend auf Cattells Konzept der objektiven Persönlichkeitsdiagnostik sowie auf aktuellen Theorien der Selbstwertorientierung wurden sechs verhaltens- und szenariobasierte Subtests konstruiert, um zustimmungsorientierte und prinzipienorientierte Selbstwertregulationstendenzen zu erfassen. Eine Stichprobe von $N = 89$ Personen nahm an der Untersuchung teil, von denen $N = 82$ vollständige Subtestdaten für die Strukturanalysen bereitstellten. Die interne Struktur wurde auf Subtestebene mittels explorativer und konfirmatorischer Faktorenanalysen untersucht. Die postulierte Zweifaktorenstruktur wurde nicht bestätigt, und die strukturelle Differenzierung zwischen den angenommenen Dimensionen blieb begrenzt. Die innere Konsistenz war über die Subtests hinweg überwiegend niedrig. Analysen zur konvergenten Validität zeigten, dass der zustimmungsorientierte Selbstwertregulationstest positiv mit einem etablierten Maß für Selbstwertregulation durch externe Zustimmung zusammenhing, während sich für den prinzipienorientierten Test kein erwarteter Zusammenhang mit einem Maß für prinzipienbasierte Selbstwertregulation zeigte. Zusammenhänge mit weiteren konzeptuell verwandten Persönlichkeitsmerkmalen erwiesen sich als uneindeutig, während divergente Validität durch geringe, teilweise nicht-signifikante Korrelationen mit einem Instrument zur Erfassung allgemeiner kognitiver Fähigkeiten gegeben war. Insgesamt spiegeln die Befunde die bekannten psychometrischen Herausforderungen der Entwicklung objektiver Persönlichkeitstests wider und verdeutlichen, dass eine kumulative Weiterentwicklung über mehrere Studien hinweg erforderlich ist, um verhaltensbasierte Indikatoren der Selbstwertregulation weiter zu untersuchen.