Proceedings of the 4th International Conference on Global Politics and Socio-Humanities
DOI: 10.54254/2753-7048/2026.34372

Measurement of Non-Cognitive Skills and Their Impact

Tianran Fu[1*], Zhengyuemeng Lei[2]

[1]*International Education, Henan University of Technology, Zhengzhou, China*
[2]*Education, University of York, York, UK*
**Corresponding Author. Email: 231170500102@stu.haut.edu.cn*

Abstract. As research on human capital shifts from cognitive abilities to non-cognitive skills, the role of these skills in education and the labour market has become increasingly prominent, making accurate measurement of them of great significance. This paper examines the measurement of non-cognitive skills by employing a literature review and comparative analysis to systematically assess major measurement approaches, including self-report scales, behavioural measures, observer reports, and digital techniques. Furthermore, issues such as construct bias and social desirability bias are prevalent; these errors can lead to attenuation bias in empirical analyses, thereby systematically underestimating the true impact of non-cognitive skills on wages and employment. Furthermore, in the Chinese context, Western scales suffer from a lack of cultural adaptation. Consequently, this paper proposes a 'triangulation' framework centred on multi-method integration, which combines localised scales, contextualised behavioural experiments and digital footprint analysis to enhance the accuracy and interpretability of measurements. The research indicates that the measurement approach itself directly influences research conclusions and policy judgments.

Keywords: Non-cognitive skills, measurement methods, measurement error, integration of multiple methods

1. Introduction

As one key dimension of human capital, non-cognitive abilities are more and more appearing to be a key variable that explains differences of individual economic attainment. The traditional theory of human capital has since long put focus on cognitive abilities that are measured by scores of standardised tests; but, recent study shows that, non-cognitive abilities (such as carefulness, emotion steadiness, anti-pressure ability and social-emotion ability) still have independent and strong forecast ability about education achievement, occupation selection, job matching, salary levels and work stability. This change has caused research to go from 'what people know' to 'how people do actions and fit themselves to their surrounding environment', hence expanding the range of human capital.

Inside the labour market, non-cognitive abilities have a particularly key function. Current research shows that non-cognitive abilities not only directly affect one person's obtain-work possibility, work performance and income levels, but also form their long-term career development

© 2026 The Authors. This is an open access article distributed under the terms of the Creative Commons Attribution License 4.0 (https://creativecommons.org/licenses/by/4.0/).

DOI: 10.54254/2753-7048/2026.34372

paths through affecting education selections and job matching procedures. Under the Chinese background, the related studies have further shown that non-cognitive abilities have an important function in explaining the gender salary gap and income differences among low-skill groups. Truly, under some specific situations, the marginal influence of them is not less important than that of cognitive abilities. This indicates that if one does not carry out the measurement and assessment of non-cognitive skills, it may bring about a systematic underestimation of the levels of human capital, hence hence it will cause biased judgments about how the labour market operates.

But, even though the importance of non-cognitive abilities is generally acknowledged, the problem of their measuring still is a key bottleneck that blocks the deeper advancement of related research. Different from cognitive abilities, non-cognitive skills are abstract, have many dimensions and rely on contexts, hence therefore making them hard to directly measure through the use of one single standard. At present, the academic circle mainly depends on many kinds of tools to describe these abilities in an indirect way, for example the Big Five Personality Inventory, the Grit Scale and the Locus of Control Scale. But, these tools display obvious differences in construct definition, measurement accuracy, and cross-cultural application scope, which can bring systematic measurement mistakes and therefore further influence estimates of economic profits brought by non-cognitive skills in empirical study. For example, measurement instruments which are based on stable personality characteristics often have difficulty to accurately forecast behavioural performance in particular situations.

Although already existing review researches have systematically made examination of the function of non-cognitive skills from the angles of education or labour economics, certain limitations still are left. Majority of research works concentrate on proving the significance of non-cognitive abilities for educational and earning results, while comparative researches on measurement instruments are comparatively few in number. The small number of researches which already discuss measurement problems usually are still limited to talks on the psychometric level, therefore they have no systematic assessment which combines measurement instruments with experiment results coming from the labor market. Therefore, the method of measuring non-cognitive skills and their influence on the conclusions of economic research still is a topic that needs further deep exploration.

Under this background, this paper carries out a systematic combing of the use of non-cognitive skills in labour market research, with focus on measurement tools. Leaving the limits of one single psychometric framework, this paper carries out a comparison analysis of main-stream measuring tools in three aspects: content covering, structure stability and standard validity, and thus further probes the possible influence of measuring error upon actual study results. It also attempts to put forward directions for optimizing measurement which is made fit for the Chinese labour market circumstance, hence with the aim of offering a reference for more correctly evaluating the human capital worth of non-cognitive skills.

2. The conceptual definition of non-cognitive skills and multidimensional theoretical models

As an important supplement which is necessary for human capital theory, non-cognitive abilities hold an important position in explaining differences of individual performance in fields like education and labour market [1]. But, this concept does not have a unified definition, and its connotations, structure and measurement methods have great differences among different academic sub-fields. These differences come from distinct research motives, theoretical traditions and obtainable data in every field, hence hence directly cause differences in the selection and effect of

DOI: 10.54254/2753-7048/2026.34372

later measurement tools. Therefore, it is first necessary to clarify multidimensional theoretical models and grasp their conceptual emphasis in concrete application situations.

2.1. Theoretical framework and labour market relevance

Non-cognitive skills, people conceptualize that they are a multidimensional construct that contains many kinds of psychological traits and behavioural tendencies, and they are not a single unitary trait. One synthesis of current researches finds four connected dimensions: the dimension of personality trait (especially Big Five), the dimension of motivation factors (for example, grit, self-efficacy), socio-emotional ability, and economy preferences (for example, risk attitude, intertemporal choice).

In the view of labor economics, these abilities are traditionally seen as a second dimension of individual difference, which stands together with cognitive skills. In this framework, particular characteristics show strong prediction ability for labor market results. It is worth pointing out that Conscientiousness has been by many researchers proven to have the most powerful cross-domain forecast ability; its capability of forecasting academic accomplishment is on the whole in accordance with measurements of intellectual ability, and it has obvious connection with work performance and salary payments. What is of key importance is that the importance of Conscientiousness still keeps comparatively stable even when job complexity becomes higher, hence it shows that this trait has wide applicability in a great many kinds of job types.

Although research fields like pedagogy and psychology may give operational definitions to these skills via observable behaviors or experimental assignments to evaluate adaptability, labor economics usually uses them to make control for already existing individual differences. Early-stage study inside this domain frequently took out single gathered structure concepts (for example self-respect or control place) from social investigation questionnaires to act as substitute measurement indicators. Therefore, the theoretical key point in the labour market circumstance puts emphasis on characteristics which explain differences in employment steadiness, job selections, and income degrees, regarding non-cognitive skills as necessary capital that supplements cognitive capability in deciding economic achievement.

2.2. Conceptual focus in application scenarios

Inside the domain of education, the emphasis of these concepts more directly connects with the processes of learning and the behaviours of learning. Research result shows that the non-cognitive abilities which have close connection with academic achievement mainly include academic behavior (for example, going to classes on time and taking part in class), academic perseverance (for example, persistence and self-discipline), academic thought pattern (for example, a growth thought pattern), and effective study methods. Evaluations also always are carried out on the basis of these observable behaviors within the school environment.

Inside the domain of work markets, this idea puts bigger stress on characteristics that are connected with job working effect and the ability to adjust. In a broader scope, non-cognitive abilities may be comprehended as including work-connected characteristics (for example diligence and cautiousness) as well as individual strong points (for example self-belief and outgoingness). Most academic circles hold that these dimensions have a very important function in the forming of labor market results. For instance, they are frequently connected with faster growth of hourly wage and higher participation degree within the labor force.

DOI: 10.54254/2753-7048/2026.34372

3. Assessment of methods and tools for measuring non-cognitive skills in specific fields

The measurement of non-cognitive skills is a crucial link between their theoretical concepts and practical applications. In the two key domains of education and the labour market, researchers and practitioners have developed a diverse range of measurement paradigms, each with its own unique strengths, limitations and appropriate contexts.

3.1. Self-report scales

Self-report scales are currently the most widely used method of measurement, typically involving the use of questionnaires to ask individuals to rate their own behavioural tendencies and psychological characteristics. Representative tools include the Big Five Personality Inventory, the Grit Scale and the Locus of Control Scale. In the field of education, this method is widely used to assess students' learning motivation, conscientiousness and emotional regulation skills. As there is a consistent positive correlation between conscientiousness and academic performance, self-report scales are frequently employed in large-scale educational assessments. Their advantages lie in their low cost and ease of administration, making them suitable for use in schools or educational surveys.

In the labour market, self-report scales are primarily used for employee selection and career assessment. For example, companies often use personality assessments to gauge a job applicant's conscientiousness, extraversion or resilience. Whilst such measurements hold some value in predicting job performance, they are susceptible to 'malreporting' in selection contexts, whereby individuals consciously embellish their responses.

Overall, self-report scales are suitable for large-scale data collection, but their main issues lie in social desirability bias and reference point bias, which affect the validity of cross-group or cross-cultural comparisons.

3.2. Behavioral measures

In order to overcome the subjective bias inherent in self-report scales, behavioural assessment provides a relatively objective source of data by inferring an individual's non-cognitive skills through observation of their actual performance in standardised situations. Its specific applications mainly fall into two categories: firstly, experimental economic paradigms, such as the trust game, risk-taking tasks and time-discounting tasks, which are used to measure individuals' propensity to cooperate, risk preferences and patience; these methods reveal individual preferences more accurately by introducing economic incentives [2]. Secondly, cognitive-behavioural tasks, such as the Stroop Task and tracking tests, which are used to assess an individual's self-regulation, attentional control and impulse inhibition abilities, and are widely applied in cognitive psychology and the evaluation of educational interventions. Overall, the core advantage of behavioural measurement lies in the objectivity of its findings, which effectively reduces the subjective bias inherent in self-reports. However, this method also faces significant ecological validity issues, namely that the simplified task scenarios in laboratories often fail to accurately reflect the complexities of classroom or workplace environments. For example, individuals who demonstrate conservative behaviour in risk-taking tasks may exhibit strong innovative capabilities in real-world contexts; similarly, individuals who display a high degree of cooperation in laboratory trust games may perform poorly in actual team collaboration due to factors such as communication styles. This situational specificity limits, to some extent, the generalisability and interpretative power of the measurement results [2].

DOI: 10.54254/2753-7048/2026.34372

3.3. Observer reports

Assessment by external observers—such as teachers, parents, superiors or colleagues—constitutes another important method of measurement, particularly in educational and workplace contexts where it has direct practical value.

In education, teacher rating scales are commonly used tools for assessing pupils' social-emotional skills, classroom behaviour and learning attitudes. Evaluations based on teachers' long-term, multi-contextual observations often provide important information beyond students' self-assessments and are correlated with students' academic achievement. However, this method is susceptible to the halo effect, whereby a teacher's overall impression of a student (such as their academic performance) may influence their assessment of specific non-cognitive skills (such as cooperative spirit). Furthermore, differences in standards among teachers may lead to low inter-rater reliability. In the workplace, evaluations by superiors, peer reviews and 360-degree assessments are common methods of external evaluation. These methods reflect, from multiple perspectives, the skills employees demonstrate in work contexts, such as teamwork, communication skills and leadership. Multi-source assessment can, to some extent, offset the bias of a single evaluator. However, such methods cannot entirely avoid evaluator bias, such as the leniency effect, the strictness effect or the influence of interpersonal relationships [3].

3.4. Digital techniques

With technological advancements, a number of emerging methods have offered new approaches to measuring non-cognitive skills.

Digital footprint research includes conducting analysis on behaviour data that are left by persons in digital spaces for the inference of their non-cognitive abilities [4]. For instance, data like login times, study routes and the keeping on doing tasks on network study platforms can indirectly reflect one student's self-control and study desire for getting good results. In the work place, communication modes on professional social network platforms and use data from project cooperation tools can also expose employees' cooperation networks and working habits. Although this method can achieve non-interference continuous observation of behaviors, it has problems in the aspects of data private safety, algorithm deviation, and the equal property of digital behaviors compared with real-world behaviors.

Neuroimaging index: Neurotechnology tools such as functional magnetic resonance imaging (fMRI) and electroencephalography (EEG) are utilized by people to investigate the neural bases of non-cognitive abilities (for example self-control and emotion adjustment) [5]. Even though these methods have promoted the cognition of the physiological mechanisms which lie under these abilities, their use is still mostly limited to fundamental research, hence this is because of high expenses, operation complexity, and restricted ecological validity.

3.5. Comparison of psychometric characteristics and domain applicability of measurement tools

With regard to reliability and validity, self-report measuring scales usually show good internal consistency reliability, but their test-retest reliability can be influenced by undulations in a person's current state; their structure validity relies on the strictness of the process that the scale is built, meanwhile criterion-connected validity—especially forecast validity—is frequently weakened by many kinds of answer prejudices. The test-retest reliability of behaviour measurement methods can

DOI: 10.54254/2753-7048/2026.34372

be unstable because of context changes or the practice influence, hence their construct validity relies on whether the task can draw out the target construct in a relatively pure way [6]. However, in the actual practice, this is frequently difficult to completely guarantee. As a comparison, the reliability of other people's ratings is greatly dependent on if the evaluating persons have got enough training and on the uniformity of evaluation standards, while their validity has close connection with the evaluating persons' chances for making observation and the correctness of their judgements. Besides this, a core problem exists in whether measurements of non-cognitive abilities can give extra prediction ability beyond the one of cognitive abilities that is their incremented validity. A great number of researches point out that such incremental validity does indeed exist [7]. For instance, when people have made control for cognitive ability, non-cognitive skills like the trait of conscientiousness still have significant prediction for academic achievement and work performance. Table 1 gives a full-scale comparison of the core features of the primary measurement approaches.

Table 1. Measurement methods: strengths and applicability

Measurement Methods	Advantages	Limitations	Applicability in the field of education	Labour market suitability
Self-report Scales	Low cost, easy to implement and highly standardised	Social expectation bias, risk of falsification	High risk of bias (suitable for large-scale screening and developmental assessment)	High risk of bias(suitable for selection and assessment, but precautions must be taken against falsification)
Behavioural Measures	Relatively objective and controllable situations	Low ecological validity,high situational specificity, and relatively high costs	Low ecological validity,high situational specificity, and relatively high costs	Suitable for assessing the decision-making or risk appetite required for a specific role
Observer Reports	Suitable for assessing the decision-making or risk appetite required for a specific role	Rater bias, inconsistent standards, high costs	Teacher, parent	360-degree appraisals and line manager evaluations are the mainstay of the practice
Digital Techniques	Objective, sustainable and unobtrusive	Ethical and privacy concerns, complex explanations, and immature applications	in the exploratory phase	In the exploratory phase

In actual operation, many kinds of measuring methods must be carefully chosen or combined, the using situation, and the special demands about reliability and validity. The future path consists in cultivating tasks that have bigger ecological validity, using technology to optimize the observation of behaviors, and utilizing multi-method triangulation to draw the map of individual persons' non-cognitive skills more accurately and fairly.

4. Potential impact of measurement error on empirical research conclusions

As abstract composition parts of human capital, the measuring mistake of non-cognitive abilities brings a systemic difficulty to cause conclusion drawing in experience studies. Because measurement tools frequently have difficulty in accurately grasping their theoretical constructs, hence the statistical bias that comes from this can greatly twist the cognition of the real influence of

DOI: 10.54254/2753-7048/2026.34372

non-cognitive skills. This chapter will from the perspective of econometrics carry out systematic analysis on the main origins of measurement error, their statistical results, and the corresponding methods for reduction.

4.1. Primary sources of error

The measurement errors on non-cognitive skills mainly come from three respects: construct validity, systematic bias, and reliability problems. These mistakes are mutually connected and together hence exert influence on data quality.

The mistake of construct validity means that the tools for measurement cannot completely capture the core content of theoretical constructs. Inside the field of non-cognitive ability aspects, this kind of mistake is especially outstanding. For instance, the scales which Western scholars have developed, such as Big Five Personality Inventory, which are rooted in individualistic cultural environments, may not when they are used in Chinese collectivist situations equivalently measure the identical psychological characteristics. This cross-cultural application deviation is a typical expression of structural validity mistake. Furthermore, different researches use various methods to measure the same non-cognitive skill construction, which brings about results that are hard to compare or integrate, hence further enlarging validity worries.

Systematic mistakes mainly contain social expectation deviation and intentional false answering [8]. In self-report measuring scales, persons who answer tend to give answers that accord with social expectations instead of reflecting their own real situations, therefore leading to systematic higher estimation of characteristics such as conscientiousness and agreeableness. These mistakes are not the random noise; they are systematically connected with individual features or measurement situations and thus are categorized into the type of non-classical measurement error.

The problems of reliability come from the dependence on situations and the instability over time that existing within non-cognitive skills themselves. The measure results for the same person may have differences among different situations or time points. Many research works thus hold that cross-section type measures on non-cognitive ability often display low consistence and stability across different time points [9]. This low dependability means that one-time point measurements have much noise, hence bringing dangers to the following statistical analyses.

4.2. Statistical consequences

The influence that measurement error exerts upon regression analysis obeys specific econometric rules, and its results change along with the category of error. Classical measurement error and attenuation bias form the basic foundational frame for people to comprehend this problem [10]. The supposition that the real model is given by Y = βX* + ϵ, in which X *stands for the real non-cognitive skill. Nevertheless, the thing is the variable X that is polluted by error: X = X * + u. Therefore it supposes the real model is set as Y = βX* + ϵ, where X* stands for the real non-cognitive skill. Under this circumstance, employing the observed variable X in the ordinary least squares (OLS)regression method can obtain an estimated value that is not a consistent estimation quantity of the true coefficient β. Its probability limit (plim)—the value that the estimator converges to in probability—is given as follows:

$$\operatorname{plim}\widehat{\beta} = \beta \cdot \frac{\sigma_{X^*}^2}{\sigma_{X^*}^2 + \sigma_u^2} = \lambda\beta \tag{1}$$

DOI: 10.54254/2753-7048/2026.34372

Here, λ is called the reliability ratio($0 < \lambda < 1$), it represents the ratio that signal variance accounts for in total variance. On account of $\lambda < 1$,the OLS estimation possesses a systematic bias in the direction of zero—this is a phenomenon that people call attenuation bias. The size of the deviation is:$\text{plim} - \beta = -(1 - \lambda)\beta$

This point makes clear that if the real return coefficient β of non-cognitive abilities is positive, OLS estimation method will therefore underestimate the influence that it brings. Take for example, in one assumed situation, if the real effect $\beta = 0.15$ and the reliability ratio $\lambda = 0.7$, the OLS estimation will become weaker to around 0.105, which has a 30% degree of underestimation. This kind of bias directly causes researchers to make lower estimation on the economic repayment of non-cognitive abilities.

The complicated disturbance from the non-traditional measuring mistake. In the actual situation, the measure error of non-cognitive ability often cannot satisfy the 'classical' assumption. For instance, the error u can have connection with the true skill level X*(for example, people who have high skill are more easy to underrate their own ability), hence it can also connect with the outcome variable Y. The results of this kind of non-classical measurement errors are more complex: the direction of errors is not certain, therefore it may cause overestimation or underestimation, hence traditional correction methods (such as instrumental variable method) may become invalid. In the assessment of non-cognitive capabilities, mistakes brought by elements such as social expectation and scenario particularity frequently possess non-classical natures, hence hence making causal deduction harder.

4.3. Mitigation strategies

For solving the deviations which come from measuring mistakes, the economic study discipline provides many methods, and researchers can pick these methods according to the kind of mistakes and the situation of whether data can be obtained.

The Instrumental Variable, which is often abbreviated as IV, method is a method that has effectiveness in dealing with classical measurement error. This principle includes the finding of an instrumental variable Z that holds correlation with the true non-cognitive skill X^*, yet holds no correlation with the measurement error u and the model disturbance ϵ. Through utilizing the relation that exists between Z and X, the IV estimation method can find out the net influence of X^* on Y, therefore it can solve the problem of attenuation bias and hence get consistent estimation results. However, to search for a strong and strictly exogenous instrument is in practice very difficult, and the effectiveness of this method cannot get guarantee under the conditions of non-classical errors.

Multiple Indicators Multiple Causes (MIMIC) model makes use of a structural equation modeling frame to at the same time deal with the measurement of latent variables (like non-cognitive skills) and their causal connection relations. In one standard MIMIC model, many observed indicators (for example, items from different scales or behavioral tests) together reflect one latent variable, while many cause variables (for example, family background, educational experiences) directly affect the latent variable. One key superiority of this model lies in that it permits the finding of the underlaid "real" skill by way of many not perfect indicators, hence it partially makes correction for measurement mistake when authors estimate its connection with result variables (e.g., wages). Even so, the traditional MIMIC model in usual cases makes the assumption that cause variables themselves are measured without having error—this is one limitation that needs careful thinking when people carry out application.

The Fixed Effects model, it is especially fit for research works which make use of longitudinal panel data. Its basic form is written as $Yit = \beta Xit + \alpha i + \epsilon it$, in which αi represents the individual

DOI: 10.54254/2753-7048/2026.34372

particular fixed effect, it captures all not changing over time individual characteristics (including certain stable parts of measurement error).

Through the comparison of changes inside the same individual along with time, this model eliminates the influence of time-unchanged confused factors and time-unchanged measurement mistakes (Table 2). Many researches have proved that total calculation results of non-cognitive abilities which are built from follow-up observation data are more dependable than measurements which are taken at one single time point, and thus can better show the connections these abilities have with long-term social and economic results. However, this kind of method cannot get rid of time-changing measurement errors and may even enlarge the temporary measurement noise.

Table 2. Comparison of calculation methods

Method	Core Principle	Applicable Error Type	Advantages	Limitations
Instrumental Variable Method	Utilizes exogenous instrumental variables for identification	Classical measurement error	Can simultaneously address endogeneity; produces consistent estimates	Valid instruments are hard to find; may fail under non-classical error
MIMIC Model	Models latent variables using multiple indicators	Measurement error in indicator variables	Integrates measurement and structural modeling; enables construct validity testing	Model specification is complex; typically assumes no error in cause variables
Fixed Effects Model	Exploits within-individual variation over time	Time-invariant individual-specific errors	Controls for all time-invariant confounders; relatively clear data requirements	Requires panel data; cannot eliminate time-varying measurement error

In the research of practice, no one single method can be applied to all situations universally. The strongest method is triangle verification—using many economic calculation methods at the same time and comparing the consistence of their outcomes. Moreover, the enhancement of the quality of measurement tools from the origin—like the development of localized measuring tables and the combination of multi-modal data—is the basic road to decreasing measurement errors and lifting research credibility.

5. Directions for measurement optimization in the context of the Chinese labour market

5.1. Specificity and challenges

The Chinese work force market, formed by the overlapping influences of digital transformation and economic structure change, puts forward special requirements for non-cognitive skills—and hence, special difficulties for the measurement of them. Based on what World Economic Forum has written in its 2025 Future of Jobs Report, close to 40 percent of skills connected to work will go through a change of form. Along with artificial intelligence taking over more regular work tasks, companies' dependence on the soft abilities of staff members has had a big growth. Materials are got from Zhaopin. com's investigation indicates that 51.6% of investigated enterprises have elevated their soft skill demand when carrying out recruitment.

Under these situations, the culture adaptation ability of measuring tools thus becomes the main challenge. When authors directly put mainstream Western measurement scales (for example, the Big

DOI: 10.54254/2753-7048/2026.34372

Five Personality Inventory) into use in Chinese situation, it will meet obvious suitability deviation. Many studies point out that Western personality measuring tools frequently use a "prescriptive universalism" method, hence they may ignore culture-specific dimensions. On the opposite side, tools which are developed by local people, such as the Chinese Personality Assessment Inventory (CPAI), bring in unique dimensions—for example "interpersonal relatedness"—which have obvious prediction ability in collectivism cultures, hence they solve the empty places in Western models. This points out the need for making or deeply modifying tools which adapt to culture, not depending on direct translate.

The demands for soft abilities in Chinese work market have the special characteristics. To speak concretely, the requirement for self-management abilities has quickly become stronger, displaying the largest speed of growth among all soft abilities. At the same time, the ability of managers has also obtained obvious growth. On a more fundamental level, cooperative coordination abilities (for instance, distant cooperation, cross-group interaction) are obtaining increasing importance, whose goal is to change individual persons from "independent performers" to "group connection persons", so as to adapt to the wide popularization of distant work and mixed group management modes.

But, the current measurement methods are facing many difficult problems: firstly, there exists an absence of clear and unified selection standards, with more than 39% of enterprises regarding this as the key hardship in identifying "super performers"; second, traditional evaluation approaches (e.g., one-on-one watch) have difficulty to suit new situations such as distant work; Third, soft abilities in their own nature are hard for direct measurement, hence this makes the appraisal procedure become complicated.

5.2. Multidimensional integration system: triangulation approach

For solving the above-mentioned problems, this paper puts forward a multi-dimensional integration system which is built on the triangulation principle. This system combines together self-report scales, behaviour experiments, and digital footprint analysis, for promoting the ecological validity, objectivity, and prediction accuracy of measurement. Can leverage behavioural data generated on localized digital platforms such as WeChat and DingTalk to assess remote collaboration and self-management abilities.

Optimization directions of self-report scales: On the basis of local tools which are like the CPAI, it is required that specialized scales be developed for specific industries and occupational fields in China (for example, manufacturing which is undergoing digital transformation), to measure emerging competence items such as "human-machine cooperation adaptive capacity." The making of scales must abide by cross-cultural adaptation guiding principles and go through strict local verification.

Design principles for behaviour experiments: Experimental designs must closely match organization situations in China, therefore to promote ecological validity. For instance, carrying out simulation of "inter-department resource coordination" works can measure a person's capability to break through departmental barriers, or the arranging of "collective decision making conferences" can assess one's influence and cooperative mode in a team. Hence, these context-included designs can more authentically mirror job-related abilities.

The innovative usage analyses of digital trail: China's very high digitized working surroundings offer abundant data resources. Through the analysis of objective digital traces—for example, response timeliness, communication network centrality, and task follow-up modes—in collaborative office software, dynamic and quantifiable outlines of individuals' collaboration, conscientiousness,

DOI: 10.54254/2753-7048/2026.34372

and other characteristics can be constructed. This method is especially appropriate for evaluating non-cognitive skill exhibition in long-distance work environments.

5.3. Policy implications and prospects

A through optimization completed measurement system for non-cognitive skills has very important application value for pushing high-quality development of China's labour market, mainly lies in the following three aspects:

The Reform of Vocational Education Evaluation: Traditional vocational education evaluations put emphasis on hard skills, but they ignore soft skills. This newly-built measurement system can support reform work through utilizing behavioral experiments to assess students' problem-solving and team cooperation abilities in real or simulated work positions, and utilizing digital traces to monitor students' development paths in the period of internships. This allows the dynamic supervision of capability growth and hence supports the formulation of individual non-cognitive ability cultivation routes on the basis of multi-dimensional evaluations.

Strengthening Recruitment Impartiality: Optimized measuring instruments can assist in constructing more impartial recruitment procedures. The standardization behavior experiments (for example, situation pressure tests) give every candidate the same platform for showing their own abilities. The objective behaviour data which comes from digital footprint analysis is able to decrease subjective deviation and the effect of social expectation in resume screening and self-report. At the same time, the measuring tools which have been adjusted according to culture can make the recognition of various kinds of talents become more fair and correct.

Dynamic Surveillance throughout Entire Career Cycles: Non-cognitive capabilities demand continuous notice and cultivation during the whole career of a single person. One monitoring system which is based on digital footprints can enable early warnings (for example, finding signs of work burnout through changes in communication modes), personal development arrangement (giving recommendations of training or promotion chances on the basis of skill situation files), and organization matching degree evaluation (doing analysis on consistency between individual and organizational cultures). This herein provides forceful sustentation for the lifelong management of talented personnel.

For pushing forward these application works, the below policy suggestions are put forward: Combine the resources of academic circles, industrial circles and government circles to build measurement standards and frameworks which suit the national situation of China; Set up special money funds to give support to the researching, developing, and verifying of measuring tools with localization and contextualization; When encouraging the innovation of data, people need to establish ethical norms and privacy protection mechanisms for the utilization of workplace behaviour data; Push the deep amalgamation of optimized measurement tools into vocational education evaluation systems and enterprise talent management systems, therefore to effectively serve the practical demand for human capital development.

Through constructing a measurement ecological system which has cultural adaptation, methodological variety and application direction, not only the assessment and cultivation of non-cognitive abilities—which is a key composition of human capital—can be carried out in a more scientific way, hence it can also supply a solid measurement basis for China's transformation from demographic dividend to talent dividend.

DOI: 10.54254/2753-7048/2026.34372

6. Conclusion

This research, in a systematic way, builds up the central function of non-cognitive abilities as one core component of manpower capital. These capabilities prove their strong forecast ability for academic achievement, job market outcome, and social adjustment, which do not rely on cognitive abilities. But, the implicit and context-reliant character of these skills brings big measurement problems, hence therefore unavoidably brings multidimensional mistakes via self-report scales and behavioral measurement methods. These mistakes, which come from not enough structure validity, systematic prejudices such as social expectation desirability, and reliability problems connected with evaluators, therefore are not only accidental disturbance. The analysis based on economic measurement methods has verified that the classical measurement mistake can cause the deviation of reduction, and it systematically underrates the actual economic repayment that non-cognitive abilities bring. Furthermore, non-classical mistakes which have correlation with family or evaluator features can make the direction of causal inference deviation not certain, hence potentially twisting policy assessments. On the aspect of academic contribution, this research puts forward a paradigm change in human capital theory from a traditional "only cognitive" focus to a combined angle of view of "cognitive and non-cognitive equal status." From the method angle, this change requests the strict usage of economic measurement tools like Instrumental Variable (IV) methods, Structural Equation Modeling (SEM), and Fixed Effects (FE) models. From the angle of policy, the results show that education systems ought to go past the single seeking of standardized test marks to bring non-cognitive ability nurturing into curriculum making. At the same time, the labor market ought to develop fairer, more aim-pointed evaluation instruments to promote the quality of recruitment and realize the effective distribution of public resources. There still exist big difficulties, which include not enough cross-culture validity, a lack of long-time direction data, and restricted ecosystem validity. The direction of future research is pointed to interdisciplinary integration and innovation which is driven by technology, for example, combining physiological index that come from neuroeconomics with big data analysis to realize dynamic measurement without intrusion, and building a "triangulation" measurement system which is on basis of Chinese social characteristics to better evaluate the long-term effects of intervention.

Authors contribution

All the authors contributed equally and their names were listed in alphabetical order.

References

[1] Kautz, T., Heckman, J. J., Diris, R., Weel, B. ter, & Borghans, L. (2014). Fostering and Measuring Skills: Improving Cognitive and Non-cognitive Skills to Promote Lifetime Success. OECD Education Working Papers, 110, 0_1. https: //doi.org/10.1787/5jxsr7vr78f7-en

[2] Humphries, J. E., & Kosse, F. (2017). On the interpretation of non-cognitive skills – What is being measured and why it matters. Journal of Economic Behavior & Organization, 136, 174–185. https: //doi.org/10.1016/j.jebo.2017.02.001

[3] Brown, A., Inceoglu, I., & Lin, Y. (2017). Preventing Rater Biases in 360-Degree Feedback by Forcing Choice. Organizational Research Methods, 20(1), 121–148. https: //doi.org/10.1177/1094428116668036

[4] Yang, M., Adomavicius, G., Burtch, G., & Ren, Y. (2018). Mind the gap: Accounting for measurement error and misclassification in variables generated via data mining. Information Systems Research, 29(1), 4–24. https: //doi.org/10.1287/isre.2017.0727

[5] Velligan, D. I., Rubin, M., Fredrick, M. M., Mintz, J., Nuechterlein, K. H., Schooler, N. R., Jaeger, J., Peters, N. M., Buller, R., Marder, S. R., & Dube, S. (2012). The cultural adaptability of intermediate measures of functional outcome in schizophrenia. Schizophrenia Bulletin, 38(3), 630–641. https: //doi.org/10.1093/schbul/sbq136

DOI: 10.54254/2753-7048/2026.34372

[6] Zhou, K. (2017). Non-cognitive skills: Potential candidates for global measurement. European Journal of Education, 52(4), 423–437. https: //doi.org/10.1111/ejed.12241
[7] Cheung, F. M. (2004). Use of Western and indigenously developed personality tests in Asia. Applied Psychology: An International Review, 53(2), 173–191. https: //doi.org/10.1111/j.1464-0597.2004.00167.x
[8] Cheung, F. M., Cheung, S. F., Leung, K., Ward, C., & Leung, F. (2007). The Chinese Personality Assessment Inventory as a culturally relevant personality measure in applied settings. Social and Personality Psychology Compass, 1(1), 465–480. https: //doi.org/10.1111/j.1751-9004.2007.00045.x
[9] Gonzalez, J. M., Rubin, M., Fredrick, M. M., & Velligan, D. I. (2013). A qualitative assessment of cross-cultural adaptation of intermediate measures for schizophrenia in multisite international studies. Psychiatry Research, 206(2-3), 166–172. https: //doi.org/10.1016/j.psychres.2012.10.015
[10] Zhou, M. and Zhang, J. (2020). Personality Psychology in China. In The Wiley Encyclopedia of Personality and Individual Differences (eds B.J. Carducci, C.S. Nave, J.S. Mio and R.E. Riggio). https: //doi.org/10.1002/9781119547181.ch317