CRITERIA FOR ASSESSING THE SCIENTIFIC AND METHODOLOGICAL POTENTIAL OF TEACHER-RESEARCHERS

Subhonova Anora Akmal qizi
Doctor of Philosophy (PhD)
in Pedagogical Sciences
English integrated course №1 department
Uzbekistan state world languages university
subkhanovaanosha@gmail.com
ORCID: 0009-0006-6461-875X
https://doi.org/10.5281/zenodo.23212080

Abstract.

Efforts to modernize the research capacity of teacher-educators depend on the ability to assess it validly. Existing assessment practices, however, tend to rely on publication counts and formal qualifications, which capture outputs rather than the underlying and developable capacity of teacher-researchers. This article develops a theoretically grounded system of criteria, indicators and levels for assessing the scientific and methodological potential of teacher-researchers, aligned with international standards of responsible research assessment. The study employs a qualitative theoretical design combining an integrative review of research on competence assessment and research capacity with a document analysis of international researcher-development and research-assessment frameworks, including the European Framework for Research Careers, the Vitae Researcher Development Framework, ResearchComp, DORA, the Leiden Manifesto, the Hong Kong Principles and the CoARA Agreement. Criteria were derived through a construct-centred approach to validity. Six assessment principles, six criteria with 24 indicators, four developmental levels with descriptors, and a multi-method assessment matrix are proposed. The criteria are motivational-agentive, epistemological-cognitive, methodological-technological, reflexive-metacognitive, axiological-ethical and communicative-impact. An integral index and a profile-based interpretation are suggested, together with a validation agenda. The article provides a comprehensive, theory-based assessment framework that moves beyond output metrics and can serve as a basis for diagnostic instruments, professional development planning and institutional quality assurance in teacher education.

Keywords: scientific and methodological potential; teacher-researchers; assessment criteria; research competence; responsible research assessment; indicators; teacher education

1. Introduction

In many higher education systems, teacher-educators are expected not only to teach future teachers but also to produce research that improves educational practice (Cochran-Smith, 2005; Tack & Vanderlinde, 2014). In systems undergoing rapid reform, such as Uzbekistan, this expectation has been reinforced by national strategies that set targets for publications in internationally indexed journals, the proportion of staff with scientific degrees and positions in international rankings (Decree No. UP-5847, 2019). Achieving these goals requires a systematic modernization of the scientific and methodological potential of teacher-researchers: the motivational, epistemological, methodological, reflexive, ethical and communicative resources that enable them to generate valid pedagogical knowledge, apply it in teaching and disseminate it to academic and professional communities.

Any effort to develop this potential depends on the ability to assess it. Assessment is needed to diagnose individual development needs, design professional development programmes, evaluate their effectiveness and inform institutional quality assurance. Yet the assessment of researchers is itself undergoing a profound international reconsideration. The San Francisco Declaration on Research Assessment (DORA, 2012), the Leiden Manifesto (Hicks et al., 2015), the Metric Tide report (Wilsdon et al., 2015), the Hong Kong Principles (Moher et al., 2020) and the Agreement on Reforming Research Assessment (CoARA, 2022) all warn against reliance on journal-based metrics and publication counts and call for qualitative judgement supported by responsible indicators.

Assessment practices for teacher-researchers in post-Soviet systems often follow the opposite tendency. Evaluation typically relies on formal indicators such as degrees, academic titles and the number of publications, which describe past outputs rather than the capacity for future development. Such indicators are also poorly suited to teacher-researchers, whose contribution includes the translation of research into teaching, curricula and school practice (Boyer, 1990). At the same time, the research literature on competence assessment offers well-developed theoretical tools (Blömeke et al., 2015; Shavelson, 2010), but these have rarely been applied to the research capacity of teacher-educators.

This article addresses this gap by developing a theoretically grounded system for assessing the scientific and methodological potential of teacher-researchers. It is guided by three research questions:

- **RQ1:** What principles should govern the assessment of the scientific and methodological potential of teacher-researchers?
- **RQ2:** What criteria, indicators and levels can represent this potential validly?
- **RQ3:** Which assessment methods and procedures are appropriate for each criterion?

2. Literature Review

2.1 Assessing competence and potential

Contemporary competence theory treats competence as a continuum linking underlying dispositions, such as knowledge, beliefs and motivation, with situation-specific skills and observable performance (Blömeke et al., 2015). Shavelson (2010) argues that competence should be measured through performance in realistic tasks rather than through knowledge tests alone. Miller's (1990) pyramid, originally developed in medical education, distinguishes four levels of assessment, from "knows" and "knows how" to "shows how" and "does", each requiring different methods. These models imply that no single method can capture a complex capacity; assessment must combine measures of dispositions, performance and real-world practice.

Assessing potential adds a developmental dimension. Unlike competence, which emphasizes demonstrated performance, potential includes latent resources that can be actualized under favourable conditions. Its assessment must therefore consider not only the current level but also the direction of development and the conditions that support or hinder it, consistent with Vygotsky's (1978) notion of the zone of proximal development.

2.2 Assessing research competence

Several instruments have been developed to assess research competence. The Research Self-Efficacy Scale (Bieschke et al., 1996) measures confidence in performing research tasks, and research on social science students has modelled research competence as a multidimensional construct including skills in formulating questions, designing studies, analysing data and communicating results (Böttcher & Thiel, 2018; Gess et al., 2017). Research on teacher-educators has focused on their researcherly disposition and research engagement (Tack & Vanderlinde, 2014; Borg, 2010). These studies provide

valuable dimensions but rarely include ethical, reflexive or impact-related components or the institutional context of research.

2.3 Responsible research assessment

International frameworks for responsible research assessment share several principles: quantitative indicators should support rather than replace expert judgement; assessment should consider the full range of research contributions, including teaching, mentoring and societal impact; and the assessment of integrity, openness and transparency should be explicit (Hicks et al., 2015; Moher et al., 2020; CoARA, 2022). These principles are particularly relevant for teacher-researchers, whose most important contributions may not appear in citation databases.

2.4 Validity of assessment systems

Modern validity theory views validity as an integrated evaluative judgement about the extent to which evidence and theory support the interpretation and use of scores (Messick, 1995). Kane's (2013) argument-based approach requires that the intended interpretation and use be stated explicitly and supported by evidence. For a system of criteria, this means that validity begins with a clear definition of the construct and a transparent derivation of criteria from theory.

3. Methodology

3.1 Research design

The study adopts a qualitative theoretical design. It does not collect primary data; instead, it develops an assessment framework through theory synthesis and model construction (Jaakkola, 2020). This design is appropriate because valid empirical instruments can only be constructed after the construct and its criteria have been defined theoretically (Messick, 1995; Kane, 2013).

3.2 Sources and procedures

Two procedures were combined. First, an integrative review (Torraco, 2005) of research on competence assessment, research competence, teacher-educators' research engagement and responsible research assessment was conducted using Scopus, Web of Science and Google Scholar, supplemented by backward citation tracking. Second, a qualitative document analysis (Bowen, 2009) was carried out on seven international frameworks: the European Framework for Research Careers (European Commission, 2011), the Vitae Researcher Development Framework (Vitae, 2010), ResearchComp (European Commission, 2023), DORA (2012), the Leiden Manifesto (Hicks et al., 2015), the Hong Kong Principles (Moher et al., 2020) and the CoARA Agreement (2022).

3.3 Derivation of criteria

Criteria were derived through a construct-centred approach in four steps. First, the construct was defined using the six-component model of scientific and methodological potential, which integrates competence theory, sociocultural and activity theory, reflective practice, adult learning and agency theories. Second, each component was translated into a criterion. Third, indicators for each criterion were identified from the reviewed literature and frameworks and retained if they were (a) theoretically linked to the criterion, (b) represented in at least one international framework and (c) observable through at least one assessment method. Fourth, developmental levels were formulated by combining the developmental logic of the European Framework for Research Careers (R1–R4) with the progression from guided to autonomous and creative performance described in models of expertise development (Dreyfus & Dreyfus,

1986). Rigour was supported by a transparent derivation table linking each indicator to its theoretical and documentary sources.

4. Results

4.1 Principles of assessment

The analysis yielded six principles that should govern the assessment of scientific and methodological potential (Table 1).

Table 1. Principles for assessing scientific and methodological potential

Principle	Content	Source
Multidimensionality	All six components are assessed; no single indicator represents the whole potential	Blömeke et al. (2015); Vitae (2010)
Developmental orientation	Assessment identifies the current level and the zone of next development, not only past achievements	Vygotsky (1978); EFRC (2011)
Multi-method triangulation	Self-assessment, performance tasks, portfolio evidence and expert judgement are combined	Miller (1990); Shavelson (2010)
Responsible use of metrics	Quantitative indicators support but do not replace qualitative judgement	DORA (2012); Hicks et al. (2015)
Contextual sensitivity	Institutional conditions are assessed alongside individual potential	Hong Kong Principles (2020); CoARA (2022)
Formative purpose	Results are used primarily for development planning, not for sanctions	CoARA (2022); ESG (2015)

4.2 Criteria and indicators

Six criteria and 24 indicators were derived (Table 2). Each criterion corresponds to one component of scientific and methodological potential.

Table 2. Criteria and indicators of scientific and methodological potential

Criterion	Indicators
1. Motivational-agentive	1.1 Intrinsic research motivation; 1.2 Research self-efficacy; 1.3 Professional identity as a researcher; 1.4 Research initiative (self-initiated projects, applications, collaborations)
2. Epistemological-cognitive	2.1 Knowledge of research paradigms and philosophy of science; 2.2 Theoretical literacy in pedagogy and the discipline; 2.3 Ability to formulate significant research problems and questions; 2.4 Critical synthesis of international

Criterion	Indicators
	literature
3. Methodological-technological	3.1 Coherence of research design; 3.2 Mastery of quantitative, qualitative and mixed methods; 3.3 Validity and reliability procedures; 3.4 Use of digital research tools, data management and responsible AI use
4. Reflexive-metacognitive	4.1 Methodological reflexivity (awareness of bias and limitations); 4.2 Self-evaluation and planning of research development; 4.3 Integration of research findings into teaching; 4.4 Inquiry stance towards one's own practice
5. Axiological-ethical	5.1 Knowledge of research integrity principles; 5.2 Ethical conduct with participants (consent, confidentiality, dual roles); 5.3 Responsible authorship and publication practices; 5.4 Open-science practices
6. Communicative-impact	6.1 Academic writing and argumentation; 6.2 Quality and visibility of publications; 6.3 Participation in research networks and international collaboration; 6.4 Transfer of research to educational practice and policy

Three features of the indicator system deserve emphasis. First, indicators 4.3 and 6.4 capture the specific contribution of teacher-researchers, namely the integration of research with teaching and school practice, which general researcher frameworks address only partially. Second, indicator 6.2 refers to the quality and visibility of publications rather than their number, in line with DORA and the Leiden Manifesto. Third, the axiological-ethical criterion makes integrity an explicit object of assessment, as recommended by the Hong Kong Principles.

4.3 Levels and descriptors

Four developmental levels were formulated (Table 3). The levels describe qualitatively different modes of research activity rather than points on a single scale, and they are not tied to academic titles: an experienced holder of a doctoral degree may be at the adaptive level when working in a new methodological paradigm.

Table 3. Developmental levels and generic descriptors

Level	Generic descriptor	Illustrative manifestations	EFRC reference
Reproductive	Performs research tasks under guidance, using established methods; reflexivity is limited	Descriptive studies; reliance on supervisors; publications mainly in local outlets	R1
Adaptive	Selects and adapts methods independently in familiar contexts; begins to reflect on limitations	Empirical studies with standard designs; publications in national peer-reviewed journals	R2
Productive	Designs original, rigorous studies; integrates research and teaching systematically; supports others	International publications; leadership of projects; mentoring	R3

Level	Generic descriptor	Illustrative manifestations	EFRC reference
		of novice researchers	
Creative-transformative	Develops new theoretical or methodological approaches; shapes research culture and policy	Influential contributions; research schools and networks; policy impact	R4

For each indicator, level-specific descriptors can be written following this generic pattern. For example, for indicator 3.1 (coherence of research design), the reproductive level corresponds to reproducing a design suggested by a supervisor; the adaptive level to choosing an appropriate standard design for a given question; the productive level to justifying and combining designs for complex questions; and the creative-transformative level to developing new designs or methodological approaches adopted by others.

4.4 Assessment methods

Following Miller's (1990) pyramid and the principle of triangulation, each criterion should be assessed with at least two methods that address different levels of evidence (Table 4).

Table 4. Assessment methods matrix

Criterion	Self-report scales	Performance tasks	Portfolio evidence	Expert or peer judgement
Motivational-agentive	Research motivation and self-efficacy scales	—	Self-initiated projects, grant applications	Interview with mentor
Epistemological-cognitive	Self-assessment of theoretical knowledge	Critique of a research proposal; literature synthesis task	Literature reviews, theoretical chapters	Expert review of theoretical framing
Methodological-technological	Self-assessment of methods proficiency	Designing a study for a given problem; data analysis task	Research reports, datasets, instruments	Expert review of methodology
Reflexive-metacognitive	Reflective self-evaluation	Analysis of limitations of one's own study	Reflective journal; research-informed teaching materials	Peer observation of teaching
Axiological-ethical	Integrity knowledge questionnaire	Ethical dilemma cases	Consent forms, ethics approvals, authorship statements	Review by ethics committee or mentor
Communicative-	Self-assessment	Writing and	Publications,	Peer review of

Criterion	Self-report scales	Performance tasks	Portfolio evidence	Expert or peer judgement
impact	of writing and networking	peer-review task	conference papers, collaborations, practice outputs	publications; stakeholder feedback

Self-report scales are efficient and capture motivational and self-efficacy dimensions, but they are vulnerable to bias and should not be used alone. Performance tasks provide evidence of what a researcher can do in standardized situations. Portfolio evidence documents actual practice over time. Expert and peer judgement integrates these sources and interprets them in context, as recommended by responsible assessment frameworks.

4.5 Aggregation and interpretation

Two complementary forms of interpretation are proposed. The first is an **integral index**. For each criterion, a coefficient is calculated as the ratio of the obtained score to the maximum possible score, $K_i = S_i / S_{max}$, ranging from 0 to 1. The integral index is calculated as the weighted sum of the criterion coefficients, $K = \Sigma w_i K_i$, where the weights w_i sum to 1. Weights may be equal or determined through expert procedures such as the Delphi method (Hsu & Sandford, 2007). As provisional boundaries for further empirical calibration, the following ranges can be used: $K < 0.50$, reproductive; $0.50 \leq K < 0.70$, adaptive; $0.70 \leq K < 0.85$, productive; $K \geq 0.85$, creative-transformative.

The second form is a **profile interpretation**, which presents the level of each criterion separately. Because the components of potential are interdependent but not interchangeable, the profile is more informative for development planning than a single index. In particular, the axiological-ethical criterion should function as a threshold: a low level in this criterion should not be compensated by high levels in others. The integral index is therefore best used for comparisons at group and institutional levels, while the profile should guide individual development.

4.6 Validation agenda

Following an argument-based approach to validity (Kane, 2013), the proposed system requires a sequence of validation steps before high-stakes use: (1) content validation of criteria and indicators by experts, using content validity indices (Lawshe, 1975; Polit et al., 2007); (2) development of instruments for each method and pilot testing; (3) confirmatory factor analysis to test the six-factor structure and its higher-order organization; (4) analysis of convergent validity with existing measures such as research self-efficacy scales; (5) tests of measurement invariance across institutions and career stages; and (6) analysis of consequential validity, examining whether the use of the system supports development rather than encouraging strategic behaviour.

5. Discussion

The framework developed in this article responds to a central problem in the modernization of research capacity: what is not assessed validly cannot be developed purposefully. By deriving criteria from a theoretically defined construct, the framework aims to avoid the main weakness of existing practice, in which available indicators, typically publication counts and degrees, are used as if they represented research capacity as a whole.

The framework is consistent with the international shift towards responsible research assessment (DORA, 2012; Hicks et al., 2015; CoARA, 2022). It treats bibliometric information as one source of

evidence within the communicative-impact criterion, combined with qualitative judgement and with criteria that metrics cannot capture, such as methodological reflexivity, ethical conduct and the integration of research with teaching. This is particularly important in systems where strong publication incentives have been associated with questionable publishing practices (Macháček & Srholec, 2022). By including integrity as an explicit criterion with threshold status, the framework applies the Hong Kong Principles (Moher et al., 2020) to the specific context of teacher-researchers.

The framework also extends existing models of research competence. Instruments developed for students (Böttcher & Thiel, 2018; Gess et al., 2017) focus mainly on cognitive and methodological skills, while research on teacher-educators emphasizes disposition and engagement (Tack & Vanderlinde, 2014). The proposed system integrates these perspectives and adds reflexive, ethical and impact-related criteria, as well as an explicit developmental structure. Its emphasis on research-informed teaching reflects the long-standing argument that the scholarship of teaching and application are legitimate forms of scholarship (Boyer, 1990), and the meta-analytic finding that research and teaching quality are not automatically linked (Hattie & Marsh, 1996), so their integration must be assessed and cultivated deliberately.

From a practical perspective, the framework can be used at three levels. At the individual level, profiles can support self-assessment and individual development plans. At the programme level, the criteria can define the outcomes of professional development programmes and serve as pre-test and post-test measures for evaluating them. At the institutional level, aggregated results combined with an assessment of research conditions can inform strategies and quality assurance in line with ESG Standard 1.5, which makes institutions responsible for the development of their teaching staff.

Several risks must be acknowledged. Any assessment system can encourage strategic behaviour if it is used for high-stakes decisions, a risk highlighted in the literature on metrics (Wilsdon et al., 2015). The proposed formative purpose and profile interpretation are intended to reduce this risk. In addition, multi-method assessment requires time and expertise, which may limit its feasibility; a staged approach, beginning with self-assessment and portfolio evidence and adding performance tasks and expert review for development programmes, may be more realistic.

The study has limitations. As a theoretical contribution, it proposes criteria, indicators and provisional level boundaries that have not yet been tested empirically. The weights and thresholds of the integral index, in particular, should be calibrated through expert procedures and empirical data. The framework draws primarily on European and Anglo-American sources, and its applicability in other contexts requires examination. Future research should follow the validation agenda outlined above.

6. Conclusion

This article has developed a theoretically grounded system for assessing the scientific and methodological potential of teacher-researchers. It proposes six assessment principles, six criteria with 24 indicators, four developmental levels, a multi-method assessment matrix and two complementary forms of interpretation, together with an agenda for validation. The framework moves assessment beyond publication counts towards a multidimensional, developmental and responsible evaluation that includes the ethical, reflexive and practice-oriented dimensions specific to teacher-researchers. It is intended to provide a basis for diagnostic instruments, professional development programmes and institutional quality assurance in teacher education systems seeking to align with international standards.

REFERENCES

1. Bieschke, K. J., Bishop, R. M., & Garcia, V. L. (1996). The utility of the Research Self-Efficacy Scale. Journal of Career Assessment, 4(1), 59–75.
2. Blömeke, S., Gustafsson, J.-E., & Shavelson, R. J. (2015). Beyond dichotomies: Competence viewed as a continuum. Zeitschrift für Psychologie, 223(1), 3–13.
3. Borg, S. (2010). Language teacher research engagement. Language Teaching, 43(4), 391–429.
4. Böttcher, F., & Thiel, F. (2018). Evaluating research-oriented teaching: A new instrument to assess university students' research competences (R-Comp). Higher Education, 75(1), 91–110.
5. Bowen, G. A. (2009). Document analysis as a qualitative research method. Qualitative Research Journal, 9(2), 27–40.
6. Boyer, E. L. (1990). Scholarship reconsidered: Priorities of the professoriate. Carnegie Foundation for the Advancement of Teaching.
7. Cochran-Smith, M. (2005). Teacher educators as researchers: Multiple perspectives. Teaching and Teacher Education, 21(2), 219–225.
8. CoARA. (2022). Agreement on reforming research assessment. Coalition for Advancing Research Assessment.
9. Decree of the President of the Republic of Uzbekistan No. UP-5847. (2019, October 8). On approval of the Concept for the development of the higher education system of the Republic of Uzbekistan until 2030.
10. DORA. (2012). San Francisco declaration on research assessment.
11. Dreyfus, H. L., & Dreyfus, S. E. (1986). Mind over machine: The power of human intuition and expertise in the era of the computer. Free Press.
12. ENQA, ESU, EUA, EURASHE, et al. (2015). Standards and guidelines for quality assurance in the European Higher Education Area (ESG). Brussels.
13. European Commission. (2011). Towards a European framework for research careers. Directorate-General for Research and Innovation.
14. European Commission. (2023). ResearchComp: The European competence framework for researchers. Publications Office of the European Union.
15. Gess, C., Wessels, I., & Blömeke, S. (2017). Domain-specificity of research competencies in the social sciences: Evidence from differential item functioning. Journal for Educational Research Online, 9(2), 11–36.
16. Hattie, J., & Marsh, H. W. (1996). The relationship between research and teaching: A meta-analysis. Review of Educational Research, 66(4), 507–542.
17. Hicks, D., Wouters, P., Waltman, L., de Rijcke, S., & Rafols, I. (2015). Bibliometrics: The Leiden Manifesto for research metrics. Nature, 520(7548), 429–431.
18. Hsu, C.-C., & Sandford, B. A. (2007). The Delphi technique: Making sense of consensus. Practical Assessment, Research & Evaluation, 12(10), 1–8.
19. Jaakkola, E. (2020). Designing conceptual articles: Four approaches. AMS Review, 10(1–2), 18–26.
20. Kane, M. T. (2013). Validating the interpretations and uses of test scores. Journal of Educational Measurement, 50(1), 1–73.
21. Lawshe, C. H. (1975). A quantitative approach to content validity. Personnel Psychology, 28(4), 563–575.
22. Macháček, V., & Srholec, M. (2022). Predatory publishing in Scopus: Evidence on cross-country differences. Quantitative Science Studies, 3(3), 859–887.

23. Messick, S. (1995). Validity of psychological assessment: Validation of inferences from persons' responses and performances as scientific inquiry into score meaning. American Psychologist, 50(9), 741–749.
24. Miller, G. E. (1990). The assessment of clinical skills/competence/performance. Academic Medicine, 65(9), S63–S67.
25. Moher, D., Bouter, L., Kleinert, S., et al. (2020). The Hong Kong Principles for assessing researchers: Fostering research integrity. PLoS Biology, 18(7), e3000737.
26. Polit, D. F., Beck, C. T., & Owen, S. V. (2007). Is the CVI an acceptable indicator of content validity? Appraisal and recommendations. Research in Nursing & Health, 30(4), 459–467.
27. Shavelson, R. J. (2010). On the measurement of competency. Empirical Research in Vocational Education and Training, 2(1), 41–63.
28. Tack, H., & Vanderlinde, R. (2014). Teacher educators' professional development: Towards a typology of teacher educators' researcherly disposition. British Journal of Educational Studies, 62(3), 297–315.
29. Torraco, R. J. (2005). Writing integrative literature reviews: Guidelines and examples. Human Resource Development Review, 4(3), 356–367.
30. Vitae. (2010). Researcher development framework. Careers Research and Advisory Centre (CRAC).
31. Vygotsky, L. S. (1978). Mind in society: The development of higher psychological processes. Harvard University Press.
32. Wilsdon, J., Allen, L., Belfiore, E., et al. (2015). The metric tide: Report of the independent review of the role of metrics in research assessment and management. HEFCE.