

PSYCHOLOGICAL DEVELOPMENT
ПСИХОЛОГИЯ РАЗВИТИЯ

Научная статья | Original paper

Artificial intelligence as a cultural tool: Heuristic potentials and limitations of the metaphor

D.E. Nikolaev ✉

Autonomous Non-Profit Organization “Research Center ‘Analitik’”, Moscow, Russian Federation

✉ dnpsy2019@gmail.com

Abstract

Context and relevance. The large-scale adoption of artificial intelligence systems in education, professional activity, and everyday life challenges established psychological conceptions of development, mediation, and subjectivity. Within cultural-historical psychology, AI is most often analyzed through the metaphor of a cultural tool, which appears to preserve theoretical continuity but may obscure the fundamentally novel nature of digital sociotechnical systems. **Objective.** Through theoretical analysis, to identify the advantages and disadvantages of considering artificial intelligence as a cultural tool within the logic of cultural-historical psychology. **Hypothesis.** The cultural tool metaphor has significant heuristic potential for analyzing the mediation of activity involving AI; however, its explanatory capabilities are limited and, when applied uncritically, lead to conceptual reductions. **Theoretical framework.** The study is based on a reconstruction of key tenets of the cultural-historical and activity traditions, as well as on interdisciplinary work in the philosophy of technology, the anthropology of digital technologies, and critical AI studies. **Results.** It is shown that AI radically breaks the link between the mediation of activity and the mediation of development, redistributes agency and responsibility, and also generates effects of reverse mediation, influencing the subject's way of thinking and self-evaluation. It is revealed that the language of the cultural tool carries a latent normative effect, contributing to the naturalization of technological change. **Conclusions.** Artificial intelligence represents the first mass cultural means introduced without a cultural-psychological project of subject development. This positions AI not merely as an object of analysis but as a challenge to cultural-historical psychology itself, necessitating a reconsideration of key concepts.

Keywords: cultural-historical psychology; artificial intelligence; cultural tool; subject development; mediation; subjectivity; digital culture

For citation: Nikolaev, D.E. (2026). Artificial Intelligence as a cultural tool: Heuristic potentials and limitations of the metaphor. *Cultural-Historical Psychology*, 22(3), 6–14. <https://doi.org/10.17759/chp.2026220301>

Искусственный интеллект как культурное орудие: эвристические возможности и ограничения метафоры

Д.Е. Николаев ✉

Автономная некоммерческая организация «Исследовательский центр “Аналитик”»,
Москва, Российская Федерация
✉ dnpsy2019@gmail.com

Резюме

Контекст и актуальность. Массовое внедрение систем искусственного интеллекта (ИИ) в образование, профессиональную деятельность и повседневные практики ставит под вопрос устоявшиеся психологические представления о развитии, опосредствовании и субъектности. В культурно-исторической психологии анализ ИИ чаще всего осуществляется через метафору культурного орудия, которая внешне сохраняет теоретическую преемственность, но одновременно может скрывать принципиальную новизну цифровых социотехнических систем. **Цель.** В результате теоретического анализа выявить достоинства и недостатки рассмотрения искусственного интеллекта в качестве культурного орудия в логике культурно-исторической психологии. **Гипотеза.** Метафора культурного орудия обладает значительным эвристическим потенциалом для анализа опосредствования деятельности с участием ИИ, однако ее объяснительные возможности ограничены и при некритическом применении приводят к концептуальным редукциям. **Теоретическая основа.** Исследование опирается на реконструкцию ключевых положений культурно-исторической и деятельностной традиций, а также на междисциплинарные работы в области философии техники, антропологии цифровых технологий и критических исследований искусственного интеллекта. **Результаты.** Показано, что ИИ радикально разрывает связь между опосредствованием деятельности и опосредствованием развития, перераспределяет субъектность и ответственность, а также формирует эффекты обратного опосредствования, влияя на образ мышления и самооценку субъекта. Выявлено, что язык культурного орудия обладает скрытым нормативным эффектом, способствуя натурализации технологических изменений. **Выводы.** Искусственный интеллект является первым массовым культурным средством, внедряемым без культурно-психологического проекта развития субъекта. Это превращает ИИ не только в объект анализа, но и в вызов для самой культурно-исторической психологии, требующей пересмотра ключевых понятий.

Ключевые слова: культурно-историческая психология, искусственный интеллект, культурное орудие, развитие субъекта, опосредствование, субъектность, цифровая культура

Для цитирования: Николаев, Д.Е. (2026). Искусственный интеллект как культурное орудие: эвристические возможности и ограничения метафоры. *Культурно-историческая психология*, 22(3), 6–14. <https://doi.org/10.17759/chp.2026220301>

Introduction

The rapid spread of AI technologies across various spheres of human activity has become one of the most significant factors transforming contemporary sociocultural practices. These changes affect not only the technical ways in which tasks are performed but also forms of thinking, communication, learning, and self-regulation, making AI an increasingly important subject of interest for psychological science.

Within the cultural-historical tradition, the interpretation of AI as a cultural tool has gained popularity (Sizikova, 2023; Noskova, 2024; Soldatova, Voiskounsky, 2021). This metaphor draws on L.S. Vygotsky's fundamental proposition concerning the key role of cultural

means in the development of higher mental functions and provides conceptual continuity between classical ideas and the analysis of contemporary technologies. At first glance, this move appears quite natural; however, it is precisely this apparent naturalness that, in our view, requires problematization. Contemporary AI systems differ fundamentally from the cultural tools whose analysis provided the basis for the development of cultural-historical psychology. They do not merely extend certain capacities of the subject but also redistribute the functions of action, decision-making, and control among the human being, the technological system, and the institutional environment, thereby giving rise to the problem of the distribution of responsibility. As a result, the traditional model of tool mediation, which presupposes

an active subject mastering and appropriating a cultural tool, proves to be a substantial simplification when applied to AI. In other words, AI does not fully meet the criteria of a cultural tool in the classical cultural-historical sense; rather, it functions as a cultural means, an element of the cultural environment that participates in organizing activity and thinking while not presupposing full appropriation and voluntary control on the part of the subject. In this regard, AI appears not so much as a new object of analysis as a kind of “stress test” for cultural-historical psychology itself. Attempts to describe it reveal the limits of applicability of key concepts (mediation, internalization, and the cultural tool) and raise the question of the extent to which the theory is capable of describing sociotechnical systems operating at levels that extend beyond individual activity.

The current state of the problem is characterized, on the one hand, by the active use of the classical terminology of cultural-historical psychology in research on the digital environment and, on the other, by the absence of a systematic analysis of the theoretical consequences of such a transfer. This creates a risk both of excessively broadening the concept of the cultural tool and of reducing it conceptually.

The relevance of the study is determined by the need for a critical examination of the metaphor of AI as a cultural tool from the perspective of cultural-historical psychology. The scientific novelty of the study lies in a comprehensive analysis not only of the heuristic potential of this metaphor but also of its limitations.

The aim of the study is to identify the advantages and risks of considering artificial intelligence as a cultural tool from a cultural-historical perspective.

Research objectives

1. To analyze the classical understanding of cultural tools.
2. To assess the heuristic potential of applying this metaphor to artificial intelligence.
3. To identify the conceptual limitations and risks of this approach.

Research question: Within what limits can the metaphor of the cultural tool be productively used to analyze artificial intelligence in cultural-historical psychology?

The concept of the cultural tool

The use of the metaphor of the cultural (essentially, psychological) tool in discussions of AI by psychologists rarely becomes an object of explicit reflection in itself. As a rule, it is used implicitly as a self-evident methodological lens that makes it possible to incorporate AI into an already familiar theoretical landscape. Yet such a conceptual transfer is not self-evident. It reflects a specific way in which cultural-historical psychology responds to the emergence of new technologies: an effort to preserve conceptual continuity and interpret radically new phenomena through categories that have already been developed and assimilated.

In L.S. Vygotsky's classical framework, the cultural tool functions as a means that mediates the subject's activity and facilitates the development of higher mental functions (Vygotsky, 1982, 2005). In the activity tradition, this conception was supplemented by an analysis of activity systems in which tools, goals, and social relations form an integrated structure (Leontiev, 1983). In both cases, the tool is understood as an element that can be mastered, appropriated, and incorporated into consciously regulated activity.

Applying this model to AI makes it possible to preserve theoretical coherence while simultaneously masking the fundamental novelty of the phenomenon itself. Unlike classical cultural tools, AI constitutes a complex sociotechnical system whose functioning is determined not only by the individual actions of the user but also by algorithmic architectures, institutional regulators, and economic interests (Floridi, 2019). In this context, the tool metaphor increasingly performs a compensatory function. It can be said to compensate for a conceptual deficit that arises when encountering an object that is difficult to describe in terms of individual mediation. The critical difference between AI and previous generations of tools is that, whereas the logic of an automobile or television can in principle be understood and appropriated by the subject, even if not at the engineering level but at the level of functional causality, the core of contemporary AI systems (the deep-learning algorithms) constitutes a “black box”. Their operation is based on statistical correlations rather than on symbolic rules accessible to conscious reflection. Moreover, AI often functions as an adaptive infrastructure that unobtrusively shapes the activity environment itself, rather than as a passive instrument applied at will.

Thus, the use of the concept of the cultural tool in the analysis of AI can be regarded as an indicator of the insufficient development, within cultural-historical psychology, of concepts for analyzing the distribution of subjectivity in multilevel sociotechnical processes. While preserving the appearance of theoretical continuity, the tool metaphor simultaneously conceals the mismatch between the classical model of mediation and the actual practices of human interaction with contemporary intelligent systems. It is precisely this mismatch that requires further conceptual analysis.

In order to assess the applicability of the metaphor “AI as a cultural tool”, it should be noted that in the cultural-historical and activity traditions, a cultural tool refers not to any means of activity but only to something incorporated into a specific type of mediation that has developmental potential. In the works of L.S. Vygotsky, the cultural tool functions as an artificially created means through which a person restructures the organization of their own mental activity (Vygotsky, 1982). What is crucial here is not the mere use of an external means but the process of mastering it and its subsequent inter-

nalization. The tool initially exists in an external, social space and then gradually becomes an internal means of regulating behavior. This process presupposes an active position of the subject involved in joint activity and relies on the development of voluntariness, reflection, and conscious control.

A.N. Leontiev continued this line of reasoning by linking the developmental effect of the tool to its incorporation into an activity system with socially determined motives, goals, and modes of action (Leontiev, 1983). In this context, the tool does not “work” by itself; it becomes psychologically meaningful insofar as the subject masters socially developed ways of acting with it. Development is understood here as a transformation of the structure of activity and of the corresponding mental functions. P.Ya. Galperin and B.D. Elkonin also contributed to the development of the concept of the cultural tool (Pavlenko, 2020).

Thus, the classical tradition makes it possible to identify several fundamental conditions under which a cultural tool can be regarded as a factor of development: active mastery, the possibility of internalization, the formation of voluntary control, incorporation into socially organized activity, and a connection with the restructuring of higher mental functions. These conditions establish stringent criteria that become problematic when the metaphor is transferred to contemporary digital technologies.

Artificial intelligence as a cultural tool: Heuristic potentials of the metaphor

Considering AI as a cultural tool has a number of significant advantages that help explain the prevalence of this metaphor in contemporary psychological discussions. The key ones are as follows.

1. The metaphor provides conceptual continuity between the classical ideas of cultural-historical psychology and the analysis of new technological forms. Understanding AI as a means of mediation makes it possible to use traditional categories of cultural-historical psychology (mediation, internalization, voluntariness, external and internal regulation, and so forth) to analyze how algorithmic systems become incorporated into human activity (Vygotsky, 2005; Wertsch, 1998).

2. The cultural tool metaphor directs research toward concrete practices of its use. Within the cultural-historical tradition, AI does not possess independent psychological subjectivity but functions as a means incorporated into the subject's activity and mediating the subject's relations with the task and the social environment. This makes it possible to analyze the influence of AI as a result of specific forms of activity in which it is incorporated.

3. The metaphor has considerable potential for studying transformations of higher mental functions. From

this perspective, AI can be considered a means of external regulation of attention, memory, thinking, and decision-making, making it possible to analyze the redistribution of functions between the human being and the cultural tool (Vygotsky, 1983).

4. Describing AI as a cultural tool opens up space for normative and ethical analysis. If AI is considered a means of cultural mediation, its design and implementation become matters of moral responsibility. Its influence on subject development depends on which forms of activity and self-regulation are supported by the technology. This approach makes it possible to avoid treating AI as a neutral or exclusively technical factor (Soldatova, Ilyukhina, 2025).

5. The metaphor facilitates dialogue between cultural-historical psychology and contemporary interdisciplinary research on digital culture and technologies. Such research in sociology, anthropology, media studies, and the philosophy of technology actively examines how technologies restructure forms of thinking, communication, and social life. However, these discussions are often conducted in terms that have only a limited connection with specifically psychological problems. By retaining its own conceptual apparatus, cultural-historical psychology can participate in the analysis of current processes without losing its distinctive specificity, incorporating the discussion of digital technologies into its own theoretical framework. This creates an opportunity to participate in interdisciplinary dialogue without dissolving into it or abandoning the specific psychological perspective on subject development.

Mediation of activity and mediation of development

Despite the considerable heuristic potential of the metaphor “AI as a cultural tool”, its use within cultural-historical psychology entails a number of substantial theoretical and methodological limitations. One of the key and least examined assumptions in contemporary psychological discussions of AI is the identification of the mediation of activity with the mediation of development. Within the logic of cultural-historical psychology, such identification is fundamentally incorrect; nevertheless, it often underlies optimistic interpretations of AI as a “developmental tool”.

In the classical tradition, the mediation of activity is not in itself equivalent to development. Vygotsky emphasized that development is associated not with an increase in the efficiency of actions but with a restructuring of their psychological organization, with the emergence of new forms of voluntariness, reflection, and awareness of one's own actions (Vygotsky, 2017; Maidansky, Rubtsov, 2025). A tool may facilitate task performance, reduce workload, and accelerate the

achievement of a result without creating conditions for the formation of new higher mental functions. Moreover, under certain conditions, it may displace those forms of activity that were regarded as developmental within the classical model.

It is precisely here that AI reveals its “problematic” character as a form of cultural tool. Contemporary AI systems are designed primarily as means of optimization, since they reduce the number of operations, automate decision-making, and minimize uncertainty and errors (Brynjolfsson, McAfee, 2014). From the perspective of activity efficiency, this is undoubtedly an advantage; from the perspective of subject development, however, the consequences of such optimization are far from obvious.

If we proceed from the activity approach, development involves transforming the structure of activity and generating new psychological forms through the productive overcoming of difficulty (Leontiev, 1983). In this context, AI often functions as a means of “removing” a problem situation because it preserves the result (efficiency) while potentially eliminating or minimizing the very process of exploratory and orienting activity that serves as a source of development. AI undoubtedly transforms mental activity, but it does so differently from classical tools, often through the delegation and “closure” of cognitive functions within an opaque circuit, radically changing the conditions for their internalization.

Thus, a fundamental mismatch arises: AI may function as a powerful means of mediating activity without simultaneously being a means of mediating development. Moreover, in a range of educational, professional, and creative practices, it may contribute to the formation of dependent forms of activity in which the subject increasingly relies on an external system and increasingly masters fewer ways of acting (Carr, 2010; Soldatova, Ilyukhina, 2025; Chigarkova, Soldatova, 2026). This contradiction undermines the very idea of an automatic “developmental function” of cultural means and calls for its radical reconsideration in relation to digital technologies.

Reverse mediation: AI as a regulator of the subject's way of thinking

An even more radical challenge to cultural-historical theory is associated with the fact that AI functions not only as a means of mediating activity but also as a factor shaping the subject's conceptions of their own thinking. Traditional concepts of cultural-historical psychology prove insufficient for describing this new reality. Accordingly, in the present study, reverse mediation is proposed as a term for a process in which a means incorporated by the subject into the performance of an activity, by virtue of the characteristics of its design and

functioning and through changes in the organization of that activity, contributes to changes in the ways in which the subject organizes, evaluates, and understands their own activity and their own capabilities. The nature and direction of this change depend on the design and modes of functioning of the means. In the case of relatively simple means, the relationship between their use and changes in modes of activity is usually directly accessible to the subject's understanding and control. In the case of complex sociotechnical means such as AI, the changes that arise are generally not initiated by the subject, are not consciously recognized by them, and cannot be fully anticipated or controlled by them. Reverse mediation differs from feedback in that the process does not merely return information about the result of an action; rather, the means itself, incorporated into the activity as its mediating component, participates, often implicitly, in changing the very way in which the subject organizes, evaluates, and understands that activity. This may be described as a transition from mediation as instrumental extension to mediation as the constitution of an environment that prescribes particular ways of being, acting, and thinking.

This raises a problem for the classical idea of internalization: what exactly is being internalized — the tool as a means or the tool as a system of external rules and evaluations beyond the subject's control? If, in the classical model of cultural-historical psychology, the tool was understood as a means through which the subject extends and restructures their mental functions while retaining an active position of mastery (Vygotsky, 1982), the situation becomes more complex in the case of AI because interaction with intelligent systems constantly places human thinking in relation to machine thinking, often within a logic of comparison, evaluation, and competition. Algorithmic recommendations, automatic generation of texts, solutions, and interpretations facilitate the formation of particular conceptions of “how to think correctly”, “what constitutes a good result”, and “where the boundary of one's own capabilities lies” (Pasquale, 2015; Crawford, 2021).

Of course, any tool can structure the activity of the person using it (the objective determination of the image of the world) (Asmolov, 2023; Kudryavtsev, 2016), but in the case of complex technological systems used in education, professional activity, the social sphere, medicine, and so forth, the issue concerns a more extensive form of influence (Nikolaev, 2025). AI begins to function as a normative regulator of thinking rather than merely as a technical means. It sets the pace, style, and format of cognitive activity; creates expectations of immediate answers; and reduces tolerance for uncertainty and prolonged processes of searching for meaning. It may be possible to speak here of a specific internalization of algorithmic standards, whereby the subject begins, to some extent, to relate their own thinking to the logic of

machine-based data processing (Floridi, 2019; Rieder, 2020). A means that, according to the classical model, should be subordinated to the subject begins to set the subject's boundaries. It shapes expectations, standards, and self-evaluation. It is no longer the person who determines how and when to use the tool; rather, the system itself, through interfaces, algorithms, ratings, and automatic suggestions, gently or even quite forcefully regulates the style of thinking, the pace of work, and even one's sense of intellectual adequacy.

From a cultural-historical perspective, this means shifting the focus of analysis from individual actions to the transformation of subjectivity. AI affects not only what a person does but also who they understand themselves to be. Interaction with AI may contribute to the formation of orientations toward external regulation, reduced trust in one's own experience, and reduced reliance on forms of knowledge that cannot be formalized. Thus, reverse mediation calls into question the classical idea of internalization as an unambiguously developmental process when applied to AI and requires its critical reconsideration under conditions of digital culture.

A tool without a subject

One of the most radical consequences of analyzing AI within the logic of cultural-historical psychology is the identification of a crisis of appropriation that undermines the very possibility of treating AI as a cultural tool in the strict sense. In the classical model, a psychological tool is something that the subject can master, internalize ("grow into"), and make part of their own system of activity regulation (Vygotsky, 2017; Maidansky, Rubtsov, 2025). In other words, a tool becomes cultural not simply because it exists but because the subject can master, appropriate, and use it voluntarily. Even complex sign systems (writing, counting, and scientific concepts) presuppose the possibility of understanding their principles and consciously controlling their use. Appropriation presupposes an understanding of the principles underlying the operation of the tool, the possibility of using it purposefully, and, crucially, the possibility of refusing to use it (Vygotsky, 2005; Leontiev, 1983). It is through appropriation that a tool becomes a factor of development.

AI clearly disrupts this logic. Contemporary intelligent systems are fundamentally opaque to the user. Their functioning is based on statistical models, large-scale computation, and infrastructures that extend beyond the boundaries of individual experience and understanding (Burrell, 2016; Pasquale, 2015). An individual user does not know precisely how decisions are made, on what data the system was trained, or what objectives and constraints are embedded in its architecture. Researchers have written about the "black box of AI" (Bathae, 2018; Marcus, Teuwen, 2024). The subject interacts not

with a tool that they master but with the outputs of a system whose logic is inaccessible to them, even when the system provides an explanation of one kind or another. Consequently, this tool cannot be mastered in the classical sense, like a hammer, a language, or a mathematical method. The user interacts with the surface of the system but does not control its logic. The tool thus seems to exist without a subject capable of assuming responsibility for it and appropriating it in the cultural-historical sense.

This circumstance has fundamental implications for the question of responsibility. In the classical model, responsibility for the use of a tool rests with the subject of activity because it is the subject who makes choices and exercises control and evaluates consequences (Rubinstein, 1957). However, the implementation of AI systems substantially complicates the localization of subjectivity. By distributed agency, we mean a situation in which some processes previously carried out directly by the subject in organizing and regulating their own activity are delegated to technical means and institutional structures. Such processes include, for example, setting and transforming a task, generating and evaluating possible alternatives, selecting a mode of action, producing and evaluating decisions, monitoring their implementation, and assessing their consequences. At the same time, the distribution of functions does not mean that all components of the system possess agency or responsibility in the same sense. A technical system may participate in the performance of an activity, alter its course, and influence its outcome without becoming a subject or bearer of responsibility in either the psychological or the social-normative sense, while simultaneously contributing to the development in the subject of phenomena that D.A. Leontiev (2024) relates to disturbances of activity self-regulation, as well as reducing the variability and diversity of outcomes of forms of human experience and activity in society (Nikolaev, Asmolov, 2025; Hutson, 2026). Therefore, the distribution of agency requires a separate analysis of which participants, and to what extent, retain the possibility of choosing, controlling, verifying, and rejecting a decision and, accordingly, within what limits responsibility can be assigned to them.

Overall, in the case of AI, responsibility becomes distributed among different people and institutional actors involved in the development, implementation, regulation, and use of the system. Such distribution blurs the very idea of psychological responsibility for a particular decision or choice and makes it impossible simply to transfer classical conceptions of the subject of activity to contemporary sociotechnical systems (Latour, 2005). We are dealing with a tool that actively affects thinking but does not belong to the subject either cognitively or socially. This calls into question the very possibility of speaking of AI as a cultural tool without substantially revising the concepts of the subject and responsibility. In

other words, in the case of AI, mediation becomes multi-level, distributed, and partially anonymous. Algorithmic systems function not only as means of performing tasks but also as infrastructural elements that shape the very conditions of possibility for acting and thinking.

For cultural-historical psychology, this means recognizing that AI operates within a logic of collective and institutional mediation in which the subject is not the sole center of control. If this is the case, then the concept of the cultural tool requires either radical expansion or an acknowledgment of its limits of applicability. Otherwise, the theory risks describing a situation that no longer exists, presupposing an individual subject where action, choice, decision-making, and control are distributed within a sociotechnical system among people, technical means, and the institutional environment.

Beyond its theoretical limitations, the metaphor of “artificial intelligence as a cultural tool” carries a substantial normative context. The language of cultural development, mediation, and evolution has a powerful legitimizing effect because it tends to represent technological change as a natural and inevitable stage of historical progress (Foucault, 1980). In this context, AI begins to be perceived less as the result of specific economic, political, and engineering decisions and more as the “next step” in cultural development (Nestik, 2019). By describing AI as a cultural tool, we seem to attribute to it in advance the status of a “natural stage” of cultural evolution, with which one should learn to live and which, in general, works “for the benefit” of development. Such a description can hardly be regarded as exhaustive (Asmolov, 2024).

This is particularly important for psychology because the metaphor of a developmental tool shifts the focus from discussing why, how, and for whom a technology is introduced to the question of how to adapt more quickly to what is presumed to be inevitable (see, e.g., Soldatova, Chigarkova, Ilyukhina, 2024; Soldatova, Ilyukhina, 2025). As a result, less attention is paid to inequalities in access to technologies, cultural differences, and the possible negative effects of AI on the development of different social groups (Benjamin, 2019; Noble, 2018). Potentially deforming effects (narrowing of independent thinking, reduced reflective control, and redistribution of responsibility) are pushed to the margins of analysis or interpreted as temporary difficulties of adaptation (Asmolov, 2020). As a result, the language of development and mediation may unintentionally legitimize and naturalize technological change, presenting it as natural and inevitable even when its consequences for human beings and society remain uncertain or, moreover, questionable.

Conclusion

The analysis has shown that AI occupies a fundamentally distinctive position among contemporary cultural

means and cannot be unconditionally incorporated into the classical model of tool mediation developed within cultural-historical psychology. Unlike traditional cultural tools, AI is incorporated into the subject's activity primarily as an infrastructural and institutionally mediated technology designed outside the logic of psychological development. This circumstance radically changes the conditions under which mediation can be regarded as a factor in the development of higher mental functions.

AI does not simply continue the trajectory of previous technological tools but represents a qualitative leap. Its uniqueness lies in the combination of opacity (“black box”), adaptability to the user, infrastructural scalability, and a claim to perform cognitive functions. It is precisely this combination that transforms AI from a simple instrument into an active participant in the redistribution of psychological functions and agency. Accordingly, cultural-historical psychology faces not only the task of expanding the applicability of its concepts but also the need for their critical revision. In particular, the distinction between the mediation of activity and the mediation of development acquires fundamental importance: increased efficiency and the automation of cognitive operations do not in themselves guarantee the formation of voluntariness, reflection, and a subject position. Moreover, under conditions of delegating cognitive functions to AI, the space for development may narrow even as functional productivity increases.

In this context, AI functions not merely as an object of analysis but also as a challenge to psychological theory. It makes visible and explicit a situation in which a cultural means is used on a mass scale without a cultural-psychological project directed toward the formation of the subject. In this way, AI reveals the limits of the classical model of the cultural tool while simultaneously pointing toward directions for its further theoretical development.

Based on the findings of the study, the following conclusions can be formulated.

1. It has been shown that the metaphor of AI as a cultural tool has considerable heuristic potential for analyzing the mediation of activity; however, it cannot be automatically transferred to the analysis of processes of psychological subject development.

2. It has been established that under conditions of using AI systems, a gap emerges between the mediation of activity and the mediation of development, calling into question the assumption that any cultural means is inherently developmental.

3. It has been demonstrated that AI functions as a means that the subject cannot fully master, appropriate, or control, which problematizes the classical understanding of the cultural tool as a psychologically mastered means.

4. It has been shown that interaction with AI generates effects of reverse mediation, influencing the sub-

ject's way of thinking, self-evaluation, and conceptions of their own cognitive capabilities.

5. It has been substantiated that the use of the cultural tool metaphor in relation to AI has a latent normative effect, contributing to the naturalization of technological change and reducing critical attention to its consequences.

The novelty of the study lies in treating AI as a limiting case of a cultural means that reveals the limitations of the classical categories of cultural-historical psychology and requires their refinement. The practical

significance of the study is associated with the possibility of using its findings in the psychological assessment of educational, professional, and digital practices based on the use of AI systems.

Future research should focus on developing expanded models of mediation that take into account the socio-technical and institutional nature of contemporary digital means, as well as on analyzing the conditions under which the use of artificial intelligence can be incorporated into a meaningful cultural-psychological project of subject development.

Список источников / References

1. Асмолов, А.Г. (2023). *Психология личности. Культурно-историческое понимание развития человека* (6-е изд.). М.: Смысл.
Asmolov, A.G. (2023). *Personality Psychology. A Cultural-Historical Understanding of Human Development* (6th ed.). Moscow: Smysl. (In Russ.).
2. Асмолов, А.Г. (2024). Библия от искусственного интеллекта: нужны ли нам новые поводыри? *Образовательная политика*, 1(97), 6–7.
Asmolov, A.G. (2024). The Bible from Artificial Intelligence: Do We Need New Guides? *Educational Policy*, 1(97), 6–8. (In Russ.).
3. Асмолов, А.Г. (2020). Облачное детство: в поисках компромисса между технологизацией и гуманизацией. *Образовательная политика*, 83(3), 6–8.
Asmolov, A.G. (2020). Cloud Childhood: In Search of a Compromise Between Technologization and Humanization. *Educational Policy*, 83(3), 6–8. (In Russ.).
4. Выготский, Л.С. (1982). *Собрание сочинений: В 6 т.: Том 1. Инструментальный метод в психологии*. М.: Педагогика.
Vygotsky, L.S. (1982). *Collected Works: In 6 vols.: Vol. 1. The Instrumental Method in Psychology*. Moscow: Pedagogika. (In Russ.).
5. Выготский, Л.С. (2005). *Психология развития человека*. М.: Смысл.
Vygotsky, L.S. (2005). *Psychology of Human Development*. Moscow: Smysl. (In Russ.).
6. Выготский, Л.С. (2017). *Записные книжки. Избранное* (Е. Завершнева, Р. ван дер Веер, общ. ред.). М.: Канон+.
Vygotsky, L.S. (2017). *Notebooks. Selected Works* (E. Zavershneva, R. van der Veer, Eds.). Moscow: Kanon+. (In Russ.).
7. Кудрявцев, В.Т. (2016). Еще раз об опосредствовании в культурно-исторической психологии. *Психолого-педагогический поиск*, 3(39), 77–88.
Kudryavtsev, V.T. (2016). Once Again on Mediation in Cultural-Historical Psychology. *Psychological and Pedagogical Search*, 3(39), 77–88. (In Russ.).
8. Леонтьев, А.Н. (1983). *Избранные психологические произведения: В 2 т.: Том 2. Деятельность. Сознание. Личность*. М.: Педагогика.
Leontiev, A.N. (1983). *Selected Psychological Works: In 2 vols.: Vol. 2. Activity. Consciousness. Personality*. Moscow: Pedagogika. (In Russ.).
9. Леонтьев, Д.А. (2024). Структура и нарушения саморегуляции деятельности: точка роста культурно-исторической деятельностной психологии. *Культурно-историческая психология*, 20(3), 27–35. <https://doi.org/10.17759/chp.2024200303>
10. Леонтьев, Д.А. (2024). Structure and Distortions of the Activity Autoregulation: a Growing Edge of the Cultural Historical Activity Theory. *Cultural-Historical Psychology*, 20(3), 27–35. (In Russ.). <https://doi.org/10.17759/chp.2024200303>
10. Майданский, А.Д., Рубцов, В.В. (2025). О системе категорий культурно-исторической психологии. *Культурно-историческая психология*, 21(3), 4–14. <https://doi.org/10.17759/chp.2025000001>
10. Майданский, А.Д., Рубцов, В.В. (2025). On the System of Categories of Cultural-Historical Psychology. *Cultural-Historical Psychology*, 21(3), 4–14. (In Russ.). <https://doi.org/10.17759/chp.2025000001>
11. Нестик, Т.А. (2019). Искусственный интеллект как когнитивный протез: трансформация образов будущего. *Образовательная политика*, 4(80), 104–117.
Nestik, T.A. (2019). Artificial Intelligence as a Cognitive Prosthesis: Transformation of Images of the Future. *Educational Policy*, 4(80), 104–117. (In Russ.).
12. Николаев, Д.Е. (2025). Искусственный интеллект как антропологический проект: от редукции к становлению (Часть 1). *Образовательная политика*, 23(4), 8–18.
Nikolaev, D.E. (2025). Artificial Intelligence as an Anthropological Project: From Reduction to Becoming (Part 1). *Educational Policy*, 23(4), 8–18. (In Russ.).
13. Николаев, Д.Е., Асмолов А.Г. (2025). Радикальная человекоцентричность: потенциал человека как источник антихрупкости организации. *Вопросы психологии*, 71(5), 3–14.
Nikolaev, D.E., Asmolov, A.G. (2025). Radical human-centricity: Human potential as a source of organizational antifragility. *Issues of Psychology*, 71(5), 3–14. (In Russ.).
14. Носкова, Т.Н. (2024). Психологические и технологические изменения деятельности человека в цифровой среде. *Психология человека в образовании*, 6(3), 329–337. <https://doi.org/10.33910/2686-9527-2024-6-3-329-337>
14. Носкова, Т.Н. (2024). Psychological and Technological Changes in Human Activity in the Digital Environment. *Psychology in Education*, 6(3), 329–337. (In Russ.). <https://doi.org/10.33910/2686-9527-2024-6-3-329-337>
15. Павленко, В.Н. (2020). Понятия «орудие», «психологическое орудие», «знак» и их соотношение. *Культурно-историческая психология*, 16(1), 122–131. <https://doi.org/10.17759/chp.2020160112>
15. Павленко, В.Н. (2020). The Concepts of “Tool,” “Psychological Tool,” and “Sign” and Their Correlation. *Cultural-Historical Psychology*, 16(1), 122–131. (In Russ.). <https://doi.org/10.17759/chp.2020160112>

16. Рубцов, В.В., Зарецкий, В.К., Майданский, А.Д. (2023). Культурно-историческая психология: современное состояние и направления развития научной школы. В: *Научные подходы в современной отечественной психологии* (с. 144–169). М.: Институт психологии РАН.
17. Сизикова, Т.Э. (2023). К итогам Международного конгресса «Л.С. Выготский и А.Р. Лурья...». *Культурно-историческая психология*, 19(2), 76–79. <https://doi.org/10.17759/chp.2023190210>
18. Солдатова, Г.У., Войскунский, А.Е. (2021). Социально-когнитивная концепция цифровой социализации. *Психология. Журнал ВШЭ*, 18(3), 431–450. <https://doi.org/10.17323/1813-8918-2021-3-431-450>
19. Солдатова, Г.У., Илюхина, С.Н. (2025). Человек технологически достроенный. *Культурно-историческая психология*, 21(1), 13–23. <https://doi.org/10.17759/chp.2025210102>
20. Солдатова, Г.У., Чигарькова, С.В., Илюхина, С.Н. (2024). Метаморфозы идентичности человека достроенного. *Социальная психология и общество*, 15(4), 40–57. <https://doi.org/10.17759/sps.2024150404>
21. Чигарькова, С.В., Солдатова, Г.У. (2026). Формула цифрового благополучия личности. *Национальный психологический журнал*, 21(1), 44–53. <https://doi.org/10.11621/npj.2026.0104>
22. Bathae, Y. (2018). The Artificial Intelligence Black Box and the Failure of Intent and Causation. *Harvard Journal of Law & Technology*, 31, 889–938.
23. Benjamin, R. (2019). *Race after technology: Abolitionist tools for the new Jim code*. Cambridge, UK: Polity Press.
24. Brynjolfsson, E., & McAfee, A. (2014). *The second machine age: Work, progress, and prosperity in a time of brilliant technologies*. New York, NY: W. W. Norton.
25. Burrell, J. (2016). How the machine “thinks”: Understanding opacity in machine learning algorithms. *Big Data & Society*, 3(1), 1–12. <https://doi.org/10.1177/2053951715622512>
26. Carr, N. (2010). *The shallows: What the Internet is doing to our brains*. New York, NY: W. W. Norton.
27. Crawford, K. (2021). *Atlas of AI: Power, politics, and the planetary costs of artificial intelligence*. New Haven, CT: Yale University Press.
28. Floridi, L. (2019). *The logic of information: A theory of philosophy as conceptual design*. Oxford: Oxford University Press.
29. Foucault, M. (1980). *Power/knowledge: Selected interviews and other writings, 1972–1979*. New York, NY: Pantheon Books.
30. Hutson, M. (2026). Bland new world: Is AI making us all think the same? *Nature*, 657, 22–24. <https://doi.org/10.1038/d41586-026-02682-3>
31. Latour, B. (2005). *Reassembling the social: An introduction to actor-network-theory*. Oxford: Oxford University Press.
32. Marcus, E., & Teuwen, J. (2024). Artificial intelligence and explanation: How, why, and when to explain black boxes. *European Journal of Radiology*, 173, 111393. <https://doi.org/10.1016/j.ejrad.2024.111393>
33. Noble, S. U. (2018). *Algorithms of Oppression: How Search Engines Reinforce Racism*. New York, NY: NYU Press.
34. Pasquale, F. (2015). *The black box society: The secret algorithms that control money and information*. Cambridge, MA: Harvard University Press.
35. Rieder, B. (2020). *Engines of order: A mechanology of algorithmic techniques*. Amsterdam: Amsterdam University Press.
36. Wertsch J. V. (1998). *Mind as Action*. New York: Oxford University Press.

Information about the authors

Dmitry E. Nikolaev, Candidate of Sciences (Psychology), Deputy Director, Autonomous Non-Profit Organization “Research Center ‘Analitik’”, Moscow, Russian Federation, ORCID: <https://orcid.org/0000-0002-3385-1640>, e-mail: dnp2019@gmail.com

Информация об авторах

Дмитрий Евгеньевич Николаев, кандидат психологических наук, заместитель директора, Автономная некоммерческая организация «Исследовательский центр “Аналитик”», Москва, Российская Федерация, ORCID: <https://orcid.org/0000-0002-3385-1640>, e-mail: dnp2019@gmail.com

Conflict of interest

The author declares no conflict of interest.

Конфликт интересов

Автор заявляет об отсутствии конфликта интересов.

Поступила в редакцию 23.03.2026

Поступила после рецензирования 08.09.2026

Принята к публикации 20.09.2026

Опубликована 30.09.2026

Received 2026.03.23

Revised 2026.09.08

Accepted 2026.09.20

Published 2026.09.30