



Between Meanings and Senses-Making Spaces: Agency and Ownership Emergence Formalization from Cultural-Historical Activity Theory Position, for an AI-Friendly Model

Mohamad El Maouch¹ · Zheng Jin^{1,2}

Accepted: 19 April 2023 / Published online: 17 May 2023

© The Author(s), under exclusive licence to Springer Science+Business Media, LLC, part of Springer Nature 2023

Abstract

Sense of agency and sense of ownership are considered crucial in *autonomous systems*. However, drawbacks still exist regarding how to represent their causal origin and internal structure, either in formalized psychological models or in artificial systems. This paper considers that these drawbacks are based on the ontological and epistemological duality in mainstream psychology and AI. By shedding light on the cultural-historical activity theory (CHAT) and dialectical logic, and by building on and extending related work, this paper attempts to investigate how the noted duality affects investigating the self and “I”. And by differentiating between the space of meanings and the sense-making space, the paper introduces CHAT’s position of the causal emergence of agency and ownership by stressing the twofold transition theory being central to CHAT. Furthermore, a qualitative formalized model is introduced to represent the emergence of agency and ownership through the emergence of the contradictions-based meaning with potential employment in AI.

Keywords Self · Agency · Ownership · Cultural-Historical Activity Theory · Meanings/senses-space · Artificial intelligence

✉ Mohamad El Maouch
elmaouch.m@zznu.edu.cn

Zheng Jin
jinzheng@zznu.edu.cn

¹ Henan International Joint Laboratory of Psychological Data Science, Zhengzhou Normal University, Yingcai Street, No.6, Zhengzhou, Henan 45044, People’s Republic of China

² Department of Psychology, University of California, Davis, Davis, CA, USA

Introduction

Sense of Agency (SoA) gained considerable attention in the past years as a concept. SoA is the subjective experience of initiating and controlling own actions and affecting the external world. It is deeply interrelated with the Sense of Ownership (SoO) of own body (or body parts), thoughts, and feelings. SoA and SoO are both considered as the grounding of the sense of Self (Moore, 2016; Braun et al., 2018; Legaspi et al., 2019). Also, SoA and SoO are important in artificial intelligence (AI) research, either regarding how SoA and SoO of AI systems affect Human's SoA and SoO, or due to the role of SoA and SoO in, for example, enhancing robot's motor and cognitive abilities by allowing it to self-attribute of action-outcome effects, and to "distinguish itself from the world" (Legaspi et al., 2019, p. 85). However, the current state of AI is still lacking SoA as a state of self-awareness. Given the current status quo, and while few SoA models exist, AI does not have SoA (see Legaspi et al., 2019; Swanepoel, 2021).

By following the guidelines of dialectical logic and the cultural-historical activity theory (CHAT)-including the functional analysis, also by building on and expanding the work introduced mainly in El Maouch and Jin (2022a), this paper attempts to discuss SoA in the context of the hard problem in AI, by considering that addressing SoA is based not only on the current impasse that AI is facing but also, the impasse in philosophy and psychology in the context of the abstract-tangible, mind-body, and subjective/objective epistemic rupture. And two, this paper employs the qualitative contradictions-based model proposed by El Maouch et al. (2019a, b, c; El Maouch and Jin (2022a) with a special focus on the formalization of self-emergence since both the agency and ownership are relative to the existence of a subject, of a self that *senses*, *acts*, and *owns*. Overall, a profound social and relational view of the self is crucial for individual agency and *self-caused events*. So, understanding the agency does not only depend on investigating the external outcome of the organism's actions. In brief, the agency is about self-realization (e.g., see Stetsenko & Arieievitch, 2004; Gozli, 2019; Stetsenko, 2019). From Leontiev we quote:

The fact is that for the subject himself, perception and achievement by him of concrete goals, mastery of means and operations, of action is a method of conforming his life, satisfying and developing his material and spiritual needs, which are objectified and transformed in the motives of his activity (Leontiev, 1978, p. 91).

In the following sections, we discuss the ontological and epistemological duality, which form the ground of the noted impasse that prevents from grasping and modeling the sense of agency along other psycho-mental processes (meaning emergence, abstract learning, symbolization, and language acquisition, generalization, etc.) that still form crucial challenges in AI, by shedding light on CHAT's alternative position focusing mainly on how reality reflection is realized through *twofold transition internalization* where the experience components are synthesized in order new qualitative levels emerge through the emergence of

contradictions-based meanings. Contradictions-based meanings are the cornerstone of the mental structure and psychological plane, including psycho-mental processes (perception, socialization, abstraction, generalization, etc.). In our case, we are focusing mainly on self, “I”, as constructed structures (also as processes), and agency and ownership emergence, as aspects (and as processes) related to these structures/processes, along with a preliminary formalization that is a crucial level in AI.

The Impasse, the Duality, and the Main Axe of Argumentation Both in AI and Psychology

By extending and overcoming the arguments regarding the hard problem in AI, it is considered that AI inherited the historical crisis in psychology based on the ideal-objective, abstract-tangible, and mind-body rupture, which form the historical argumentation axis in philosophy; hence, creating epistemological and methodological difficulties especially in modelling and system formalization. The impasse is that neither modern psychology nor AI can realize the synthesis of both poles of the rupture; therefore, the impasse keeps proliferating new models by mechanically combining the existing models which belong to one of the two poles without qualitatively synthesizing them. As a result, the rupture is leading psychology, and AI, to become eclectic, quantitative, and positivist. However, the crisis in AI is more severe, and the field is even more eclectic and proliferative than psychology because the philosophical worldview in AI is not apparent and explicit as in psychology, in addition to the disintegrated nature of the task of the majority of models in AI.

The noted context expanded the AI’s margin allowing it to reproduce the high number of models residing between the top-down and the bottom-up frameworks or the hybrid framework through the mechanical combination of both on the ground of connectionism. In a nutshell, the duality and the rupture resulted in the axis of argumentation in AI about how to allow the system to reproduce semantics and to make sense, or to overcome what is called the meaning barrier in AI, due to the mainstream tendencies in the field follow the formula “all we needed was more of the same” (Dennett, 2017, p. 86), supported by the brute force of speed and computing power (Ekbia, 2008), preventing the mainstream models from being emergent and forming a purposeful and autonomous whole (see, Dreyfus, 1981; Bickhard & Terveen, 1996; Vygotsky, 1997a; Cruse et al., 2000; Haselager, 2005; Carter, 2007; Haselager & Gonzalez, 2007; Goertzen, 2008; Klochko, 2008; Dafermos, 2014; Quintino-Aires, 2016; Gjorgjioska & Tomicic, 2019; El Maouch & Jin, 2022a).

The above clarifies that the debate around SoA is exhausted in the field, even when the concept of SoA is not mentioned explicitly every time the hard problem (i.e., consciousness) was investigated, e.g., autonomous systems. However, all the related aspects noted in the literature regarding SoA were handled already since the debate about SoA falls into the framework of semantics and sense-making as the main axis of argumentation in AI (see El Maouch & Jin, 2022a), including the *reflection* problem, causal inference, and goals’ ownership, among others (Moore,

2016; Braun et al., 2018; Swanepoel, 2021). For instance, SoA requires numerous conditions including purposeful actions, a sense of self, causal structure, actions initiation, being subjective, having the first-person position, finding common sense, and a sense-making process, among others. The aforementioned reveals the same axis of argumentation in the field of AI in general whereas the emergence of meaning forms the cornerstone being the representation of grasping the internal content of the phenomena (abstraction and generalization), which means that the epistemological rupture and ontological duality are represented in every *unsolved* function in the field (see El Maouch & Jin, 2022a). By investigating the main theories and models that explain SoA, one finds the same duality noted earlier, represented either, from one hand, by the objective-naturalist or the pure sensualist pole by inferring about SoA from the neurophysiological correlate of actions and sensorimotor experience (this is the implicit or low-level/non-conceptual measures of SoA, e.g., the comparator model-CM), whereas, from another hand, by the subjective-phenomenological pole by inferring about SoA through the self-report of the subject experience through questionnaires (this is the explicit or higher-order/conceptual measures of SoA, e.g., the retrospective inference-RI through sense-making).

Although some models have been introduced, as a compromise, to overcome the disadvantages of both the implicit and the explicit, such as the multifactorial weighting model (MWM) (and later the Bayesian cue integration theory), however, here we still find the noted rupture between the subjective (idealist) and the objective (biological/naturalist) poles, because the content of the phenomena is epistemically and methodologically unreachable for mainstream methodologies and the combination of both poles is a mechanic addition as same as the hybrid models in AI (e.g., see Legaspi et al., 2019; Moore, 2016; El Maouch & Jin, 2022a). This position leads some to propose that SoA could be illusive or not even exist in the first place because what is measured is the effect and not the organic and structural emergence of SoA. So, one can conclude that the emergence of the self, as an authentic and objective crucial component in the psychological plane that plays a crucial role, is still out of reach and even out of consideration for mainstream methodologies.

Thus, being based on positivist positions where the represented models rely only on a pure sensual ground, the mainstream tendencies represent person-vacuum, depsychologized, and sub-personal standpoints, and consign consciousness to a metaphysical limbo (e.g., see Haselager, 2005; Haselager & Gonzalez, 2007; Goff, 2009; Dennett, 2017; Frankish, 2021). This is why the next paragraphs are dedicated to investigating the emergence of the self as a crucial component and to the potential formalization of self-emergence and how it is employed in AI. Doing that will be based on exploring the legacy of the cultural-historical activity theory (CHAT) and of Vygotsky, especially the *Thought and Language* text (Vygotsky, 1986) regarding the process of reflection. Moreover, modeling self-emergence is based on the attempts to formalize dialectical logic in AI, and especially in robotics (see El Maouch et al., 2019a, b, c; El Maouch & Jin, 2022a). In investigating the development of the mental plane, Dialectics is about considering the wholeness, the internal contradiction through the unity of contraries, and the continuous movement of objective reality (e.g., see Vygotsky, 1997a; Bolis & Schilbach, 2020). The interest in CHAT's epistemic and methodological positions is considered relatively new (in

the late nineties), also main parts and texts of the theory were not well known in the international academic community and the mainstream academy compared to the establishment of the theory that goes back to around a century ago by Vygotsky, his colleagues, and successors, especially the key text *The historical meaning of the crisis in psychology: a methodological investigation* (e.g., see Van der Veer & Valsiner, 1991; Vygotsky, 1997a; Stetsenko & Arievitch, 2004; Goertzen, 2008; Hyman, 2012; Roth, 2014). Therefore, CHAT is still far from being fully explored especially regarding the debate on the formation of self and self modelling, that is crucial for SoA in AI.

Activity, Mind, Dynamic System of Meanings (DSM), and Self

In this paragraph, we will expand on what is introduced in El Maouch and Jin (2022a) regarding the internal structure of activity-based reality reflection and its contradictions-base being the causal origin of mind formation, with a special focus on the self-emergence model and how AI could benefit from it; but first we will introduce in brief CHAT's guidelines in comparison to mainstream and conventional approaches.

CHAT in Brief

Unlike mainstream approaches, reality reflection and the production of thought/knowledge cannot be statistically deduced from the organs of sense alone. In addition, the qualitative phenomena cannot be reduced to the quantitative phenomena (Leontiev, 1978; Mikhailov & Darglish, 1980; Gribanov, 1981; Vygotsky, 1986, 1997a). So, the reflection of the world is “accumulated not only directly at the sensory level but also higher cognitive levels...In other words, the ‘operator’ of perception is not only simply the previously accumulated associations of sensation” (Leontiev, 1978, p. 41), as mainstream sensualism does. Additionally, the basis for higher mental functions and the developmental psychological processes cannot be explained by the mechanisms of the brain, as mainstream biologism does (Luria, 1966, 1976; Leontiev, 1978; Vygotsky, 1997a). Therefore, for the sake of our current investigation, we are based on the *philosophy of activity* instead of the philosophy of the biological body/brain, or the philosophy of mind, the philosophy of pure physical behavior. CHAT investigates the consciousness and the *mental* both phylogenetically and ontogenetically as sociohistorical phenomena. For instance, the mind (e.g., identity, intentions, personality, self, mental functions, consciousness, and the unconscious) is the outcome of the system of social activity. Also, the activity system and its outcome (the mind), in addition to other material phenomena such as nature and society, are governed by dialectical laws. Thus, the real activity (and not the phenomenological experience) is where to start (the first level of) the investigation of mind formation and mental activities (the second level, i.e., the psychological level). This is why the neurophysiological brain (the third level) is only the plane where the first two levels are represented.

The aforementioned leads us to conclude that in the process of adaptation, needs, and desires are what causally drive the active organism. In this context of being driven by needs and desires, and in seeking satisfaction, and as Hegel noted, the activity confronts the environment's resistance and the resistance of the agent's body Itself as well, i.e., the natural forces and matter's characteristics, including the readymade sociohistorical conditions, relationships, objects of desire, social actors, tools, and culture including language and symbols (Mikhailov & Daglish, 1980; Marx & Engels, 1996, 1997, 1998). It is "in this process, by acting on external nature and changing it, he [the human being] at the same time also changes his own nature and acts upon it" Vygotsky, 1997a, p. 87). The activity is "a molar, not an additive unit of the life of the physical, material subject" (Leontiev, 1978, p. 50). So, the meaning-based (psychic) reflection, as a *twofold transition*, "emphasizes the constant flow of activity as the source of mind and self" (Stetsenko & Arieivitch, 2004, p. 484). Firstly, the subjectivity must include the practical role of the organism (the reflected practical goal-oriented activity) as a component in the activity once again, hence, allowing the active components of activity to emerge (the agency of mental activity and actions). And the second transition is when the formed active components themselves become the object of another reflection, i.e., self-reflection. For example, let us consider the action of the organism pushing an object. The act of pushing, when reflected for the first time, becomes the mental component of an objective active actor. However, the second reflection of this mental component is *sensed* as an action triggered and owned by the agent themselves (*I am doing*). The reflection from the position of *CHAT* is totally different from mirrored and recurrence reflection that one can face in mainstream studies. Unlike mainstream mechanical reflection, self-recurrence became a qualitative process and gain its qualitative content only by introducing new qualities to the system. For instance, when the self is formed the mental plane gains new components, relations, and laws. The formation of the *self* initiates a new qualitative path of the agent, which can be condensed by the argument that the self is the "embodiment of a meaningful project... that reflects and also organizes the most significant aspects of one's life" (Stetsenko & Arieivitch, 2004), allowing the self to play an evaluative role regarding experience's signs, hence, form the core of interpretability and reality reflection in general. The self, as a formed component, gives the ground for new subjective needs, hence, becomes a constraint of the activity, that shapes the mental structure. Therefore, the emergence of the self, as the core of subjectivity, is an objective-material process that entirely originated in the process and the flow of activity (Stetsenko & Arieivitch, 2004). So, based on modeling subjectivity, *CHAT* includes in an interactive framework, crucial components of AI systems, e.g., needs and goals (e.g., the network constraints noted by connectionism, the internal drivers noted by embodiment direction, the rules noted by formal symbolism), interpretation, semantics, and intrinsic meanings (see e.g., Bickhard & Terveen, 1996; Haugeland, 1997; Ekbia, 2008; Dennett, 2017). However, for *CHAT*, and unlike the mechanical tendency of other mainstream approaches, these components have a qualitative and causal role by being part of the dialectical framework of contradiction formation. In general, we cannot exhaustively

explore CHAT, however, what forms our interest is how CHAT engages with the problems at hand. Specifically, by shedding light on how meanings emerge and form the core of the reflection process. Unlike the analysis by the element of the positivist method, it is by following the analysis by units, meaning emergence will be the basic unit in the whole complex system of intelligence, including self-reflection, interpretation, abstraction, generalization, emotions and intellect unification, language acquisition and concept formation (Vygotsky, 1997a). Overall, to be able to grasp and formalize the meaning emergence is to introduce the ground for the subjective-meaningful entity (self) in the context of experience (Froese, 2007). In meaning emergence lies the crucial qualitative transfer from the tangible into the abstract, which will bridge the epistemological gap noted earlier. The above shows in general how CHAT differs from mainstream approaches, and later in the text more differentiations in detail will be introduced regarding the formation of self as a process, and not as an absolute entity or a pure phenomenological state. Most of all, the mainstream methods use in *connecting* the experience's content are either reinforcement learning (RM), or analogy making. However, both RL (the trial and error method, based on behaviorism in psychology) and analogy making (based on constructivism in psychology; see Drescher, 1986, 2003) form open challenges. For instance, only by relying on “temporal sequences and by the application of a mathematically conceived formula of the functional interdependence of phenomena,” Piaget replaced the “explanation of phenomena in terms of cause and effect by a genetic analysis” (Vygotsky, 1986, p. 96); and does “not provide a ‘mathematical’ model for integrating the elements of experience or concerning how to shift from one stage of development to the next, and it is not clear how these different levels of abstraction operate (Ekbia, 2008; Stojanov, 2009; Kelley & Cassenti, 2011)” (El Maouch & Jin, 2022a). On the other hand, RL, “attempts to explain the development... from the viewpoint of the mechanistic principle of the accidental combination of heterogeneous elementary reactions” Vygotsky, 1997a, p. 201). Therefore, it excludes the thinking process by reducing development to a stimulus-response relationship (Vygotsky, 1997a). Furthermore, “the informational function of reward and punishment is limited because there is no understanding of the stimulus-response relationship” (Bedny & Karwowski, 2006, p. 350). This is the positivist standpoint which is devoid of an active person. Therefore, behaviorism (and RL) lies short in explaining the problem of interpretability and mattering due to the lack of a reference entity for meaning formation. When “reducing intelligence to merely sensor-actuator mechanistic behavior through a process of blind trial-and-error, RL cannot understand the higher complex mental activity that results in long-term learning” (El Maouch & Jin, 2022a, p.9). Moreover, for RL, semantics and active perception are open problems Vygotsky, 1997a; Cruse et al., 2000; Bedny & Karwowski, 2006; Carter, 2007; Kober et al., 2013). For Vygotsky: “The description ‘this animal is running away from some danger,’ however insufficient it may be, is yet a thousand times more characteristic for the animal’s behavior than a formula giving us the movements of all its legs with their varying speeds, the curves of breath, pulse, and so forth” Vygotsky, 1997a, p. 277).

Next, we investigate with some expansion above guidelines.

Activity and the Formation of the Mind

For CHAT, the activity is always goal-directed in the context of realizing needs and desires through chains of actions. And through the internalization process of the “active” context, the internal psychological plane is formed. Therefore, the starting point of mind development, structure, and components should always be about analyzing the objective external activity and its conditions. Here, the activity is the molar, where the given elements (needs, desires, actions, conditions, sensory inputs, and mental activity) are all synthesized. Therefore, being an internalization of the active organism’s existence in the environment, the reflection outcome is always active (e.g., the active perception). Activity is not merely the physical and tangible actions and operations.

Activity is a molar, not an additive unit of the life of the physical, material subject. In a narrower sense, that is, at the psychological level, it is a unit of life, mediated by psychic reflection, the real function of which is that it orients the subject in the objective world. In other words, activity is not a reaction and not a totality of reactions but a system that has structure, its own internal transitions and transformations, its own development (Leontiev, 1978, p.51).

Activity is a system having its internal qualitative content and abstract meanings representing the interaction among experience’s various components and their flow that also cannot be reduced to the physiological level of the brain (e.g., see Luria, 1966; Leontiev, 1978). Therefore, it is understanding the reflection process (and its role in the activity system) that allows us to fill in the gap between the subjective and objective poles of existence because the subjective is yet a specific form of objective; hence, it allows us as well to uncover the laws of transition from tangible into abstract, from physical into psychological. From Leontiev we quote: “Psychology is a concrete science dealing with the origin and development of the reflection of reality by man, which takes place in his activity and which by mediating it fulfills a real role in the activity” (Leontiev, 1978, p. 32). The outcome of internalization/interiorization through reflection is not the “representation” of isolated elements but resulted in the whole plane of consciousness with its various levels, including the formation of mental functions, self, and personality.

Reflection is not the classical one-sided theory of projection of the external world in the brain; instead, it is the reflection of the totality of the individual-world relationship as a social practice in a shared context among various social individuals. The aforementioned position contrasts the mainstream reductionist tendency of the experience to the sensorimotor level, hence (see [CHAT in Brief](#) section), losing the qualitative meanings of the experience (Leontiev, 1978). Moreover, it is always the external actions that precede the internal ones. It is in the external context where the internal/mental operations initially originated. It is when “in the course of carrying external actions with external objects subsequently continue to develop in the plan of internal mental activity according to its logical-genetic laws” (Leontiev, 1978, p. 58). For Vygotsky: “the subject of activity also is *subject to* and *subjected to* his/her activity and, therefore, shaped by the environment,” and it is “in actual human praxis,” where “the subject is changed

in both material and ideal terms” (Roth, 2014, p. 11). It is “in this process, by acting on external nature and changing it, he [the human being] at the same time also changes his own nature and acts upon it” (Vygotsky, 1997a, p. 87). Therefore, not only the mental functions but also the “player” parties on the subjective stage (e.g., self, personality) have to be analyzed based on the investigation of the objective process, and their “basis is the category of objective human activity, the analysis of its integral structure, its mediation and the forms of psychic reflection that it generates” (Leontiev, 1978, p. 112). But that does not mean that the *internal* psychological stand above the *external* one (from where it originated) and is not separated from it, but an essential twofold connection preserved with it and form one system in action and consciously developing.

Overall, the emergence of consciousness itself (the conscious human) is a historical process that belongs to a certain level of development of objective existence of people that is the social activity, i.e., the context of the shared process of production (e.g., see Leontiev, 1978; Stetsenko & Arieivitch, 2004).

Furthermore, to understand an ongoing process is to actually understand the history of that process in action (Vygotsky, 1978). Thus, since our focus is on the gain of the aspects of *ownership* and *being active (agency)*, the following paragraph will investigate the objective origin of self-formation and agency in the activity system and its internal structure.

Self as a Flow and Objectively Originated

From the *nun-reductionist ontological vision of CHAT*, Stetsenko and Arieivitch (2004) attempt to overcome the *ontological mutism* regarding the analysis and the establishment of the principal sources of the self as *important agentic dimensions*. Also, the analysis of what self is made of, where it is located, and its role. Hence, it is to overcome the mainstream dualism between subject-object and between individual-society by allowing their dialectical unity, as well as to integrate the current disconnected perspective of self, in contrast to, first, the individualistic-mentalist (cognitivist) position about subjectivity, and two, overcoming the fusion of *self with the context* that belongs to the positivist position (which embedded to a certain level in the pure naturalist or biologist paradigm). Also, it is an attempt to overcome the relatedness, responsiveness, and dialogism approaches to self.

One have to start from the argument that it is the *ever-growing* demands in practice and in social life that form the origin of the phylogenetic, and obviously, the ontogenic development and complexity of the mental and subjective plane. And according to the main insight law of Vygotsky that “all the intrasubjective processes first originate as intersubjective ones” (Stetsenko & Arieivitch, 2004, p. 483) (e.g., see Vygotsky, 1978), as another version of Marx’s statement that human essence is the totality of social relations (see Leontiev, 1978). However, that does not mean in any way that both forms of existence (subjective and objective) do not have their own qualitative path of formation, and the subjective cannot be reduced mechanically to the objective. So, according to Leontiev:

The human self has no history, and no logic of functioning and developing, beyond the history and logic of functioning and developing of human practical purposeful activity (Stetsenko & Arievidtch, 2004, p. 484).

Moreover, according to the notion of *twofold transition*, the object of activity first leads and directs the subjective activity since no activity is objectless. And secondly, the reflection of this activity - which includes the process of activity, the characteristics of object, and the natural and social-cultural (material) laws that constrain (as well as enable) the activity, including the human body characteristics such as needs and desires - will be, in turn, embedded in/absorbed by the reflected image of the object. This notion emphasizes that the source of self and mind is the constant flow of activity. This goes with Aristotle's argument of the self as an activity (e.g., see Group of contributors, 2020; Charles, 2021). In this term, the self is the crystallized product of the activity processes. The selves are the functional response to the requirement of the objective social context to an orienting and regulating element (i.e., the self) (e.g., see Leontiev, 1978; Stetsenko & Arievidtch, 2004). Again, from Leontiev we quote:

The personality of a man is in no sense preexisting in relation to his activity; just as with his consciousness, activity gives rise to personality (Leontiev, 1978, p. 105).

A brief view of schizophrenia and pandemic-related lifestyle studies reveals the role of activity flow in mental coherence and the contextual origin of self as a process. For instance, during Covid-19, the disturbance of daily activity is considered the source of psycho-mental problems (e.g., see Benke et al., 2020; Chew et al., 2020; El Maouch & Jin, 2022b; Gopal et al., 2020; Guesoum et al., 2020; Ma et al., 2020; Massad et al., 2020; Odriozola-González et al., 2020; Pandey et al., 2020; Shi et al., 2020; Singh et al., 2020; Tee et al., 2020; Wathelet et al., 2020). Also, the activity structure and time usage are considered deeply related to mental aspects and their impairment, including self-regulation, executive function, and quality of life (e.g., see Luria, 1976; Leontiev, 1978; Vygotsky, 1986, 1987; Powers et al., 2007; Hagell et al., 2012; Hodgekins et al., 2015; Meijers et al., 2015; Zaytseva et al., 2015), therefore, the “activity level in individuals at risk of schizophrenia is extremely important from a preventive point of view” (Semenova, 2020, p. 3). Also, the disturbance of life narrative, that is the flow of life experience (activities), is related to mental coherence (e.g., see Davidson, 1993; Roe & Davidson, 2005; Lysaker et al., 2010; Allé et al., 2015).

And as Munsterberg et al. (1910) noted about investigating unconsciousness, where Vygotsky (1997b) himself quoted, the topic at hand (here it is SoA and SoO) cannot be answered only by observation and studying the facts as such and needs to follow the epistemological arguments, in contrast to the current mainstream methodologies. So, “it is a philosophical problem which must be settled by principle before the explanation of the special facts begins” (Munsterberg et al., 1910, p. 22).

The Twofold Transition, Extended

The twofold transition is crucial when we extend it and apply it to the reflected image once again. We mean by that not only the mental and the environmental molded together, but also the components of the mental, that are formed via the reflection, will be molded together as well, but at the same time they keep their qualitative existence. In this way, the object of the reflection will not be the external activity only (and its content), but also the mental plane itself will be the object of the reflection. It will be reflected to itself one again. So, the twofold transition works on, one, the level of the external objects, and, two, it also works on the level of reflecting the internal (what is already reflected) mental plane, including the formed mental products and processes. So, again, not only the external processes are reflected, but also the internal processes will be reflected as well (see Fig. 1).

Self: the Objective Side of the Subjective Existence

The internal plane, which is formed through internalization, will become an object of another reflection. In this second reflection, the internal plane will be reflected (as an object of reflection) in itself again. This recursive reflection is where the self is rooted, as well as the formation of the internal dynamics of the mental plane following the objective-material laws that govern reality (see [Activity and the Formation of the Mind](#) section). Therefore, we are not talking only about an external flow, but also about an internal flow, which will be both the objects of continuous reflection.

It is not in our intention to exhaust the complexity of the outcome resulting from this multi-level continuous reflection, but for the sake of our analysis, what can be said is that the multi-layered and multi-components mind, provides the key to understand the structure of the psychological plane, as well as the constituent components,

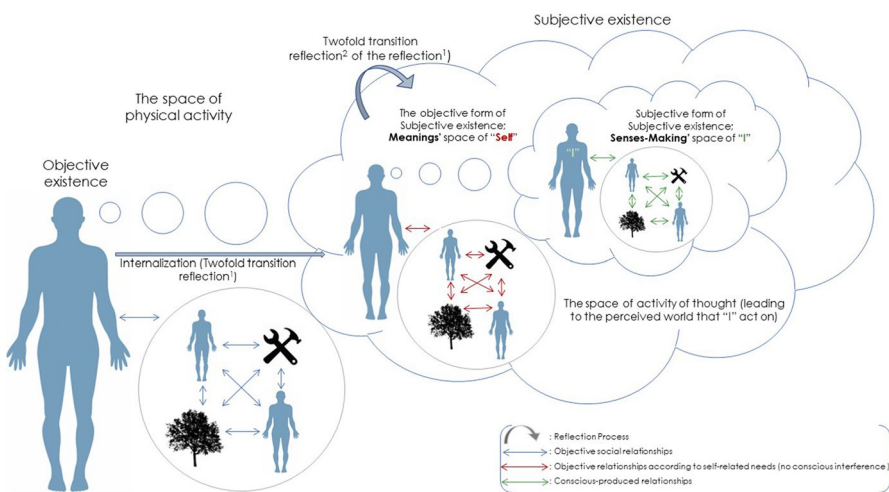


Fig. 1 Twofold transition realization through existence’s spaces

e.g., the conscious and unconscious, the “I” as representing the thinking component of the mind, and imagination.

Therefore, we will focus only on how the emergence of self, as an outcome of the continuous reflection of the internal plane/flow, that is, in turn, is a reflection/internalization of the external plane/flow in a twofold transition relationship, can be formalized and represented in an *AI-friendly* model (borrowing the concept from Swanepoel (2021)). Also, it gives us the chance to differentiate between agency and SoA, and between Ownership and SoO.

First of all, by stressing the point that the “self” (or self as object) and “I” (or self as subject) should not be confused (e.g., see Tagini & Raffone, 2010; Woźniak, 2018), we differentiate between the thinking field “I” and the field of self since both emerge from another source and fulfill different functional roles, although they may appear the same for the phenomenological position. For the current phenomenological mainstream, the *being* is equal to the *phenomenon*, and there is “no difference between what it seems and what it is” leaving us only with “phenomenology”, and according to Feuerbach, “we must differentiate thinking and the thinking of thinking” and “between the data of consciousness and the process” (Vygotsky, 1997b, p. 115). For more about the functional analysis of mental processes and mind in general in the context of CHAT one can go back to Luria (as well as Oliver Sacks) (see Proctor, 2022).

In addition to the differentiation between self and I, it is also crucial to locate the causal origin of how the formed psychomental structure/entities gain their own internal movement. Under the effect of the material laws (following the guidelines of dialectical logic), and in relation to the new subjective products, new needs, and new motives emerge, by considering that to CHAT the object (product) is the source of motives emergence. The previous may explain what is meant by that the self, as any other phenomenon, is driven by a sort of ‘self-movements’ having internal moving relations, contradictions, and mutual transitions, hence, allowing us to consider that the structure of the self is “a relatively stable configuration of principal motivational lines” (Leontiev, 1978, p. 135) (e.g., see Stetsenko & Arievitch, 2004).

In other words, the self is the product required by the praxis (*a practical relevance*) at a certain level of complexity and hierarchy. So, the self is “oriented toward real-life practical tasks and pursuits of *changing something in and about the world* (including in oneself as part of the world)” (Stetsenko & Arievitch, 2004, p. 494).

In a nutshell, the self is an original (material) product of the continuous flow of activity, being reflected/internalized to form another mental flow of its own. What is crucial to note here is that the interaction of both flows will be also reflected (on itself), forming a stable product representing the relational distinctive existence of the self’s bio-physiological holder (the human body) from the environment (bordered with the tactile sense/skin specifically). The role of tactile experience in (active) perception, body image formation (the perceived body), body ownership, and self-discrimination, are pointed out and investigated both in humans and robotics (e.g., see Schütz-Bosbach et al., 2009; Medina & Coslett, 2010; Serino & Haggard, 2010; Tsakiris, 2010; Girãoa et al., 2013; Kappasov et al., 2015; Kilteni & Ehrsson, 2017; Huynh et al., 2019).

The above is ground to understand the proposition about how the self is the subjectification of objective social relationships, connections, and ties. It is the distinct mutual existence of the individual that is at the same time a confrontational existence with the natural and social existence including the confrontation with the nature of the bio-physiological body that the individual possesses (Vygotsky, 1997a). The previous position allows us to understand what is meant by the self as a reflection of *all the things* and that all the content of the consciousness is self-related (e.g., see James, 1890; Woźniak, 2018). It is this confrontation and the disturbance of flow that always brings a given phenomenon into awareness (Vygotsky, 1987).

As such, the formation of a *distinct* self, as a subjective product of an objective *distinct* existence, provides the object of the emergence of the central psychological need for *self-realization* centralized around self-value. In that sense, self-realization is the subjective form of the totality of the sociohistorical existence of individuals or their *principal motive*. By being as such, the self-related emerged needs are the outcome of self as a general *meaningful project* of a given person and the *leading activity* that relates the individual with the historical transformative movement of a given society. The *leading activity* defines the dominant relationship with reality that leads and drives other aspects of individuals' development. Overall, the self, is the source of emerged psychological needs (being the derivations from the *universal* need that is the need of self-realization and social recognition) according to the specific sociohistorical conditions. For Hegel and Marx, as both are two main modern dialectical philosophers, history is a *process* (and subject formation is a process as CHAT revealed later on), which means that the subjective existence is the outcome of realizing:

themselves and their particular aims in actions determined by particular human needs...The subject (or the individual) is the series of its actions... each individual makes himself a member of society through a process of self-realization and self-determination; and he realizes himself as a social being in society through his activity, diligence, and skill. (Türkyılmaz, 2015, p. 252, 253).

For some individuals in certain conditions, these needs of being recognized by society are realized through scientific contribution, for others, it is to focus on their appearance (being beautiful), whereas some people engage in a political project or social movements, and for another group of people, it is about being loved, or simply, stifling and resisting change. Such a *leading activity*, that positions people to meaningfully contribute to the world, makes the self the generator of ever-developing goals and motives following the development of the ongoing transformative activity in a given sociohistorical context and specific conditions. *The aforementioned represents the agency of self from a practical standpoint* about the world (including the standpoint about own self), which makes the cultural values, including scientific and social thoughts, crucial contributors to forming the self (e.g., see Leontiev, 1978; Stetsenko & Arieivitch, 2004) (see Fig. 1).

Furthermore, although we noted earlier that the complexity resulting from the multi-level reflection will not be exhausted in this paper, however, we will only draw attention to a crucial point that is related to our analysis, especially to the interaction

between SoA and SoO. We mean the role of multi-level reflection in building various *activity spaces/planes*.

To state it from a functional analysis point of view, the self is the reflection of the outcome of the effect of the environment on the distinct objective existence, concerning the accompanying emerged needs, desires, and goals. The outcomes are the meanings forming the content of the individual's experience governed by the law of self-realization, e.g., is the individual recognized by the social context and environment or not, and the way of being recognized or not, or the interaction of the body with the environment and the social and functional outcomes of that interaction. For example, it is the case of certain physical deficiency, such as losing the ability in a specific limb or sense organ (e.g., see Stetsenko & Arievitch, 2004). It is the psychological form of self-preservation (since the bio-physiological is the other form). And by following the twofold transition, these meanings would be embedded in the structured world (see Fig. 2). Therefore, being a product resulting from the distinct existence in interaction with social relationships governed by the newly emerged needs of this product (i.e., the self), and following the twofold transition, we can state that the self is the *objective side of subjective* existence within the twofold space (which is objectively originated in the first place). In other words, according to the twofold transition, since the outcome of the internalized reflected reality will be embedded in this reality, therefore, there is always subjective in the objective, and there is always objective in the subjective. The self is an objective product (an existence) in the material reality which belongs to the psychological plane. However, the “I” originated from another functional source to serve other practical tasks.

“I”: the Subjective Side of the Subjective Existence

“I” is the product of the crystallization that occurs for the internalization of the active aspect of the individual existence in the context. “I” is rooted in the reflection into the psychological plane of the active part of the practical processes. This is why the tools used in practical activity also become the tools in the thinking process of

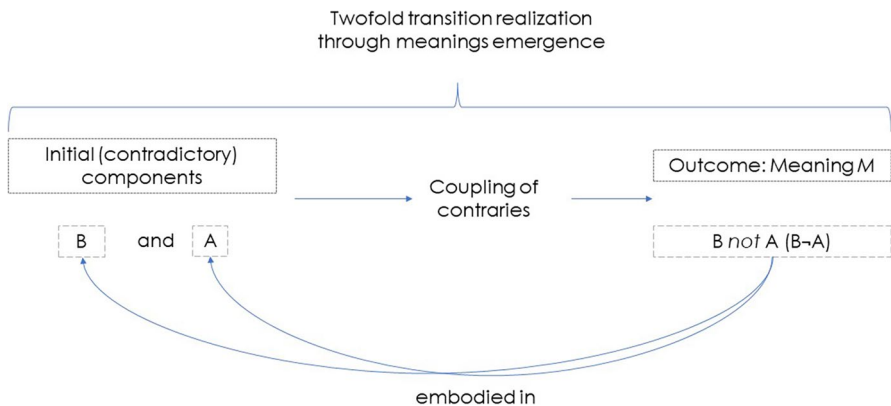


Fig. 2 Twofold transition realization through meanings emergence

“I”, such as knots and languages (e.g., see Vygotsky, 1986). The “I” deals with real objects (mental and physical), including the self as an objective object, using various tools. With some practical objects, the required tools are the knowledge (practical and scientific) regarding the objects’ characteristics and laws, e.g., the laws of physics and chemistry; while with mental objects (e.g., the self), the required tools are mostly the philosophical, psychological, and social knowledge using mainly the language as the object here requires a high level of abstraction to grasp the internal content and the laws of the psychological life itself (in relation to the sociohistorical laws at the final analysis). Therefore, “I” makes sense of the reality it deals with, and these senses do not necessarily embed the meanings included in reality phenomena. The senses that are

reflecting motives engendered by actions of life relationships of man may not adequately embody their objective meanings, and then they begin to live as if in someone else’s garments...This which makes it possible to introduce into the individual’s consciousness and impose on him distorted or fantastic representations and ideas (Leontiev, 1978, p. 93).

Thus, if the self and its needs are the sources of meanings in the mental plane, the “I” is the product that makes sense of these objective meanings using the mental tools (based on language and other signs as sociohistorical and cultural products). Here, the “sense is viewed as a result of subjective reflection of the meaning” (Bedny et al., 2014, p. 139) (please, see the discussion about the verbal and nonverbal as a point of departure in analyzing the various level in psychological structure in Vygotsky, 1986, 1997b); Hence, this is considered the space where meanings and words unite in concepts according to Vygotsky (1986)’s theory about the interaction between thought and language. This is why the narrative of the “I” is not always representing the experience of the self and the system of meanings embedded in the self’s story (e.g., see Stetsenko & Arieivitch, 2004). So, “in given circumstances, the lack of correspondence of sense and meaning in individual consciousness may take on the character of a real alienation between them, even their opposition” (Leontiev, 1978, p. 91).

Thus, “I” can be considered as the *subjective side of subjective existence* in the twofold space. For instance, this distinction between the self and the “I” is the ground of what is known as the conscious-unconscious dichotomy, especially since this dichotomy “is of decisive methodological importance for each psychological system. This problem is fundamental for our science, and its very fate depends on the way it is solved” (Vygotsky, 1997b, p. 110). Also, it is the ground of alienation of modern individuals (e.g., see Vygotsky, 1997a). Therefore, using the concepts of dialectics, the self belongs more to the space of *necessity*, while the “I” belongs to the realm of *freedom* that is governed by the realm of *necessity*. In Marx’s words, freedom lies in knowing the necessity (Marx & Engels, 1998). Therefore, by using the dialectical concepts, in the mind, what is in the space of *necessity* (the self and the formed meanings) forms the *potential* of the *freedom* space (the “I” and thinking with senses) (see Vygotsky, 1997b, p.119). The opposition between the sphere of meanings and the sphere of senses is not absolute and unhistorical; instead, it is conditioned by the ever-developing activity and living conditions (the social

existence and life relationships, including culture, as a whole) in the final analysis (see Leontiev, 1978, p. 91, 93).

The previous distinction between, from one hand, the *objective side of subjective existence* which is mainly structured from an objective system of meanings, i.e., the space where the self belongs, and from the other, the *subjective side of subjective existence* that is structured from subjective sense-making, i.e., the space where the “I” acts, we say, this distinction is the realization of CHAT denial of the subjective-objective dichotomy. Moreover, since meanings appear, in mainstream discussions, to be merged with senses it “makes it necessary in analysis to isolate the personal sense as still another forming system of individual consciousness” (Leontiev, 1978, p. 92).

Here we reach a crucial level of the investigation at hand that is exactly the differentiation between the agency and SoA, and the ownership and SoO. Based on the noted distinction between the space of meanings and the space of senses, one can differentiate between the objective content of agency and its reflection on the *phenomenological* level. So, it is on the second level (phenomenological one) where the mainstream study tries to investigate agency and ownership, only by studying their phenomenological appearance.

Later (see [Consciousness Plane, Dynamic System of Meanings \(DSM\), and Contradictions-Based Meanings Emergence](#) section), we will introduce the structure of the meanings that constitute the agency and ownership, as both are processes forming different sides of the activity, and show some implications that allow us to explain some phenomenal illusive experiences regarding both agency and ownership in some empirical studies (e.g., the hand illusion, the body transfer illusion (BTI), enfacement illusion, etc. ; see Braun et al., 2018).

Since the roots of the mental structure lie in the activity system (including social and cultural tools such as knowledge, as well as the historical conditions of activity, also, how and to which limits they develop the process of self-realization) nor the *borders* between the sense-making space and meanings space is not stable, neither the borders between the I and the self under the tension and the continuous change between the borders of *the kingdom of necessity and the kingdom of freedom* along the historical mastering of the truth about personality in on the path of mastering the truth about the society (see Vygotsky, 1997a).

Despite the different ontological and epistemological positions, similar differentiation between the self (also “Me”) as an *object*, representing the objective content of subjective existence, and the “I” as a *subject*, representing the “thinker” side and the subjective form of subjective existence, can be noted in the analysis of James (1890), Wittgenstein (1958), and others (e.g., see Woźniak, 2018). Moreover, various contradictions in the field matter of differentiating Self and “I”. For instance, one contradiction is how we can preserve a sense of mine-ness while we lose the sense of agency or how to clarify the limits of ownership (the matter of the extent of self-relatedness). These sorts of contradictions and problematics lead the mainstream tendencies to be pushed towards metaphysics by allocating the “I” as a *knower* outside the matter of study as an *ultimate entity* (e.g., see Woźniak, 2018). However, by adopting the twofold transition, these matters find their resolution (see [Twofold Transition, Agency and SoA, Ownership and SoO](#) section). So, along

with the differentiation between the “I” and the self, we also differentiate between the agency and the sense of agency, and between the ownership and the sense of ownership.

The separation between the mental spaces and components is only methodological and never epistemological because the psychological plane is a unified whole where all its components interact coherently and dynamically in a continuous flow. They exit through each other.

Twofold Transition, Agency and SoA, Ownership and SoO

For now, we already introduced how the self and “I” emerge from distinct functional roles in the context of the activity following the internalization process by reflecting the external into the internal, and later, the internal into the internal. Also, it is noted that according to the twofold transition, both the activity and the activity materials (e.g., the external objects) will be merged and unified. However, it is crucial to add that according to the twofold transition, there is also another twofold relationship between the *subjective subjectivity*’s space of the “I” (representing the active component) and the *objective subjectivity*’s space of the self (representing the crystallized distinctive objective existence concerning the material law of self-realization). Therefore, not only the external objects (including the body) of experience will have the aspect of ownership (as a distinctive aspect of self-related), but also the thoughts which are generated by the thinking “I”. Also, the agency is not only an aspect of the active “I” but also becomes an aspect of the objects of the activity, including the external objects as well as the internal objects (especially the self). To say, each phenomenon, let it be physical or mental, enters the psychological plane it will be injected with agency and ownership because it enters that plane through the system as an object or a tool of activity.

Therefore, both the ownership and the agency are original aspects of the reflected components, including material/tangible objects, people, and thoughts. This is why on the phenomenological level of “I” the illusions appear. This is the ground where we can explain why everything on the phenomenological level can be sensed with a certain level of ownership and agency because, in fact, they are all synthesized according to the twofold transition, and there is a presence of each component in the others (see Fig. 1). Therefore, each object of activity, let it be mental or physical, will be embodied by both the agency and ownership. For instance, the tunnel, the hammer, the door, and the car, all represent in themselves the activity they were part of (or the potential ones), which means they are being embodied with the ownership aspect, as well as the agency aspect. All are the extension of the self, by being part of the internalized twofold transition, as well as they are an objectification of the active “I”. So, the objects not only embody *what can I do with these objects*, but also embodies *my needs and goals*, hence, embody *me fulfilling my needs and desires by using them*, so they are always as *artefacts in use* and “no longer part of the environment but part or extension of the actor’s body” (Albrechtsen et al., 2001, p. 20) as *functional organ*, because the agency is originated from the objects of the activity and not

in isolation from them. The objects are important in *shaping the use*. And these *functional organs* can be external (e.g., scissors) as well as internal (e.g., action plans and mental models) (see Leontyev, 1981; Albrechtsen et al., 2001). Thus, the “affordances are inseparable from the original operational conditions, and that these affordances have become part of new functional organ” (Albrechtsen et al., 2001, p. 20).

Actually, there is only a mutual existence and not a parallel one. This is why we have certain illusions. For example, although objects have distinctive existence from the individual, the sense of ownership of certain objects arises from the reality that all the objects of activity are synthesized -according to the twofold transition- with the self when they are reflected-internalized in the psychological plane, allowing to that all internalized objects have a certain *trace* of self. For instance, we consider that this is ground to understand the self-relatedness or “Inclusion of Other in Self” scale (see Aron et al., 1992; Woźniak, 2018). Also, the illusion of the multi-agentive “I” (see Woźniak, 2018), makes the experience of some thoughts lacks of agency simply because the agency itself is attached to a certain active player in the psychological field, which is in turn the outcome of the real activity flow. Therefore, a severe difference in the context and the conditions of activity (e.g., the domain of activity, and the social roles/identity related to this domain) leads to the differences among the active players, i.e., different “I”(s), which is the root of such differences in the phenomenal experience. So, individuals can objectively have various agentive “I” although the degree of difference among them is not that high, or easy to be noticed by the individual themselves. This relative difference in agentive “I” is the objective ground of the phenomenal difference in the sense of ownership since some objects and phenomena objectively belong to a “different” actor (in the same psychological plane). This is a clear example of how the “I” and the self are not absolute entities but originated in the context of the activity flow. The previous is another form of distortion when the halted activity can be the source of mind disintegration in the case of schizophrenia or schizotypal personality (e.g., see Semenova, 2020).

To summarize, the agency is an original aspect related to the active role of the “I” in the context of solving practical problems in relation to the agents’ needs, desires, and goals, whereas ownership is also an original aspect related to the distinctive objective existence (as a distinctive body) regarding the sense of self-related phenomena. However, this distinction does not mean that they (agency and ownership) are totally distinct but coexist due to the unified existence of the psychological plane governed by the effect of the twofold transition that synthesizes the activity levels. This is why both agency and ownership can be sensed on the phenomenological level clearly or with certain deviations due to their coexistence according to the effect of the twofold transition. So, the weight and the importance of certain objects can be higher than others which means that the presence of a sense of ownership and sense of agency will differ following the aspects of the activity system they are part of (some people, places, and objects are considered more self-related than others due to their meanings in the activity system). In this sense, we can understand the following:

Thus, the power of the object of feeling is the power of feeling itself; the power of the object of reason is the power of reason itself; and the power of the object of will is the power of the will itself (Feuerbach, 1989, p.5).

The aforementioned position allows us to understand Feuerbach (1989)'s argument that in religion god is the reflection of self and represents "the highest subjectivity of man abstracted from himself" (Feuerbach, 1989, p. 31) because the self *floods* all over the sense-making space, hence, the "I" will *sense* the self everywhere in that space (see Fig. 1). This allows us not only to understand Feuerbach's statement but also enables us to analyze the psychological bases of other philosophers' ideas by investigating the origins of these ideas in the interaction between meanings-space and senses-making spaces. That might be another position to achieve what Lakoff and Johnson (1999) attempted to present as a cognitive investigation of philosophical ideas, or embodied philosophy. And what we are calling here the meanings-space goes with what they called the *metaphors*.

Furthermore, this distinction between the meanings space and sense-making space, also the distinctive origin of agency and ownership allows us to understand the complexity and confusion usually raised around the topic. For instance, the self has its phenomenal version, but also the "I" has its phenomenal version as well, since both are objective/authentic and distinct components in the psychological plane. Moreover, agency and ownership have their authentic existence as processes (in the meanings space), but they also have their phenomenal existence (in the sense-making space). We consider that such twofold existence of each component/process is the origin of the phenomenal illusions, as well as the origin of the confusion in the debate around the topic, for instance, by attributing a *metaphysical* nature to subjectivity (e.g., see Braun et al., 2018; Woźniak, 2018).

Due to the twofold transition, the agency, the ownership, the object, and the operations (how the actions are executed in certain conditions) are all synthesized. Now, we reach another crucial step in the investigation regarding the internal structure of synthesizing the existence's components.

Consciousness Plane, Dynamic System of Meanings (DSM), and Contradictions-Based Meanings Emergence

Going further in our attempt to formalize the SoA in relation to self as a meaningful project requires us to analyze in more detail the role of meanings in CHAT which is how the reflection process is realized, and it is through the internal structure of meanings we can move into the level of formalization the processes under investigation (i.e., agency and ownership) (for an expansive view, please see El Maouch & Jin, 2022a).

For CHAT, the meanings play a decisive role in forming the fabric of consciousness, including the regulation of thinking, and other processes such as perception and agent-environment interaction, not only for higher mental functions but also for most straightforward actions through the Dynamic System of Meanings (DSM) where the experience components are synthesized and integrated

(e.g., actions, sensory inputs, needs, goals, desires, and emotions). Similar views about the crucial role of meaning are noted by others from different theoretical directions, e.g., the meaning generator system (MGS), and the role of the meaning in the functional production of knowledge, agent-environment interaction, and perception (see Menant, 2003, 2015).

The formation of meanings is a:

Special form of psychological reflection” and “was defined as the procedural and structural development of personal meanings in the course of human activity, which integrated the processes of creation and the functioning of the cognitive structures (images, concepts, and knowledge), goals, and the emotional and motivational components of thinking (Babaeva et al., 2013, p. 12).

This is also supported by considering the brain as a unified system of functions that could be only understood based on the psychological processes and not vice versa. The brain comports interacted analyzers structured in a specific hierarchy that relates the sense organs with the cortical layers through various levels of analyzers. These interacted analyzers are the neurophysiological version of DSM and its realization on the brain level. Moreover, the consideration of the self and the mental plane as a continuous flow of activity is explained by the literature through the concept of *the working brain*. The brain cannot be understood only in its working (active) state (Luria, 1966, 1976).

The above explains why the meanings, sense-making, semantics, and meaningful experience form the axis of argumentation both in psychology and AI, not only for the sense of agency but also for various crucial mental functions (e.g., perception, language acquisition, and concept formation). However, following the connectionist, quantitative, and statistical tendencies in understanding the formation of meaning, the mainstream models did not yet grasp the internal structure of meaning formation.

In El Maouch and Jin (2022a), the guidelines of dialectical logic and CHAT propositions are followed to present a formal model of meaning emergence based on contradictions. For dialectic, one crucial law of objective existence, including nature, society, and psychological level, is the internal contradictions that are the source of the development of a given phenomenon. Something becomes active and gains its self-drives only due to its internal contradictions.

From a dialectical perspective, change and development are a result of contradictions between events occurring in different progressions, such as biological, psychological, or cultural sociological progressions. The resolutions of these contradictions, or crises, provide the basis for further development - both positive or negative - of the individual... (Riegel, 1979, p. x).

The contradictions-based development is the formal representation of various crucial aspects of the self, such as *being through becoming*, in addition to the transcendental aspect of any developmental system which contrasts the homeostatic or autopoietic view regarding the adaptive systems (Kosok, 1976; Klochko,

2008). Both *being through becoming* and being as *transcendental* reflect the active aspect of self, and the psychological plane in general, as a process of transformation. It is by changing reality, we will change ourselves (e.g., see Vygotsky, 1997a; Stetsenko & Arievitch, 2004).

For dialectics, the contradictions are the source of the qualitative shift from one level of existence to another and where new products appear to existence. So, the moments of contradictions are the representation of the qualitative content of a given phenomenon. It is by grasping the moment of contradiction, through the negation relationship, one can grasp the abstraction (and later the generalization) of a given situation by grasping its internal content through isolating and mentally retaining the relationships among things and determining their functional role, as crucial for functional reflection and causal inference, away from the positivist and quantitative-statistical tendencies (e.g., see Kosok, 1966; Tikhomirov, 1988; Davydov, 1990; Klochko, 2008).

In a nutshell, grasping the moment of contradiction and its structure allows the grasping of meaning, and making sense of the given situation. Thus, the contradictions-based meaning is the cornerstone in understanding the functional reflection process, and how the transition from tangible to abstract, from objective to subjective, is made. Moreover, the crucial negation is crucial for subjective judgment (e.g., see Smith et al., 1995). To Hegel, power is gained by facing the negative and not turning the face away from it. Tarrying “with the negative is the magical power that converts it into being” (Miller, 1977, p. 19), similar to the argument about the role of pain in AI (Carter, 2007; Dennett, 2017). For Vygotsky (1997a), the negative is the productive side. The previous allows us to understand the position about the *error-based genuine* in AI, where the system passes through variation creating new criteria and signals (e.g., see Bickhard & Terveen, 1996). Indeed, meanings are the process “where a person is solving a problem and the detection of contradictions in the objective properties of the task,” and “the attempt to adjust the contradictory properties of the object leads to different representations of the primary operational meaning of the solution attempt” (Babaeva et al., 2013, p. 13). It is Vygotsky (1986) who said that development is facing the perturbation and overcoming them in the context of fulfilling own needs and desires. From a similar position, Menant (2011, 2015, 2020) considered that meanings are the outcome of the interaction of internal constraints with the environment. Even the debate about SoA stressed the importance of the disruption of flow and the intention-action-outcome chain, and how AI should infer the disruption’s causal origin (Legaspi et al., 2019), which we consider as an indirect representation of perturbation/problem and the functional content of the phenomenon noted earlier.

The role of contradiction is usually noted in AI; however, it was mistreated by separating the two contradictories (the two poles of contradiction), in contrast to the dialectical unity of contradictories. Also for some, and against dialectical view, one of the two contradictories is considered the real absence of the other; on the contrary, one contradiction (the antithesis B) is always the positive absence of the other (thesis A). The Hegelian structure of contradiction-based development supports our investigation since Vygotsky himself was “engaged by the Hegelian formula ‘thesis, antithesis, synthesis’...” (Vygotsky, 1986, p. xii).

Various attempts to formalize dialectical logic basic arguments are noted (Counet, 2012). However, for now, the main attempt to focus on in our preliminary formalization attempt is allowing the system to relate components via the negation relationship *not* (\neg) to provide the ground for meaning emergence that is a contradiction-based component, whereas coupling two contradictories is $B \text{ not } A$ ($B \neg A$) (see El Maouch et al., 2019a, b, c; El Maouch & Jin, 2022a).

And since agency and ownership are, like any other mental phenomenon, meanings-based processes, such coupling of the contraries are the formalized internal content of agency and ownership structure.

Formalization of Agency and Ownership

Although some attempts noted the necessity of dialectic in investigating and formalization agency, they kept the parallelistic, quantitative, and connectionist epistemological and methodological position (e.g., see Bolis & Schilbach, 2020). By differentiating between the agency and the ownership as processes that belong to the space of meanings, and SoA and SoO as phenomena that belong to the space of senses, it is crucial to state that we will narrow our goal to formalize the agency and ownership, not the SoA and SoO. What is important for now is the function of an AI-friendly model of agency and ownership, and not how a model makes sense of itself. Although sense-making provides the system with qualitative functions, as we discussed earlier (see [The Twofold Transition, Extended](#) and [Twofold Transition, Agency and SoA, Ownership and SoO](#) sections), the decomposition in the psychological plane between the meanings space and the sense-making space is due to the *disproportionately* of how though grasp (make sense) the objective internal content of the experience (meanings). Therefore, in the case of AI systems, we could eliminate this disproportionality because we are allowing the system to grasp the meanings directly, due to our pre-study (and pre-design) of these meanings. Therefore, following our theoretical and epistemological position (in this case it is the guidelines of CHAT and dialectical logic as a result of the historical accumulation of knowledge that can be transferred to the system we are designing) we will *lend* our sense-making space to the system, by supposing that our sense-making space is based (as science should be) on revealing the objective relationships of reality (see Leontiev (1978)'s position about the misleading of sense-making in [“I”: the Subjective Side of the Subjective Existence](#) section). Another component that is latent in the model is the “I”. As discussed earlier, “I” is the psychological representation of the active player of the activity. However, in the artificial system, this component is latent since the system is already designed to execute a specific function. Although the “I” can be explicit in designing the system, hence, increasing the reflective abilities and the transparency of that system, however, for the sake of keeping our focus on the emergence of agency and ownership, we will stick with space of meanings, and keep the latent presence of the “I” for now in designing the system.

In the following paragraphs, we will introduce how both agency and ownership will emerge based on the emergence of meanings as contradictions-based entities/structures.

Meanings as Synthesis and Realization of Twofold Transition

According to the twofold transition, the experience components and outcomes are synthesized. Such synthesis is achieved in the construction of meaning as the *main cell* in the formation of the psychological plane by being the representation of abstraction and generalization through grasping the content of the phenomenon (see [Consciousness Plane, Dynamic System of Meanings \(DSM\), and Contradictions-Based Meanings Emergence](#) section). Also, the internal structure of meaning formation is based on coupling contraries (entities) to form a contradictory moment. In the language of the organism, and following self-realization - as a leading activity (see [Self as a Flow and Objectively Originated](#) section) - the biological and psychological ever-developing needs will face the forces of the environment (including the body's forces) to be realized. In this context of realization attempt, perturbations will exist. Hence, contradictions will rise between the desired state of the agents and the current state regarding the given conditions. Let us consider that in a given moment the need for food is triggered, while the agent in this current state cannot reach the source of food. So, a perturbation exists. This perturbation is the contradiction between the desired state DS^i and the current state CS^i . The desired state includes the agent's need for food N^i , while the current state includes the agent's current actions A^i , and the surrounding conditions Cd^i (sensory inputs through the agent's senses). So, the internal structure of the contradiction Ct^i will be $CS^i \text{ not } DS^i$, or $CS^i \rightarrow DS^i$. The coupled contraries indirectly tie the experience's content to construct a meaning M^i through the negation relationship *not* or \neg as a crucial relationship in representing the internal content of the given situation from the agent's *subjective position/point of view*. The meaning M^i represents the experience "I need food". Furthermore, according to Kosok (1976), the outcome of the contradiction (in this case it is M^i) will be embodied in the initial components of the contradiction, i.e., DS^i and CS^i , which means that the needs and the sensory-motor content will embody the meaning M^i .

Now, let us suppose that, in a given current condition Cd^j , Ct^i (the perturbation of not satisfying the need for food) is solved and the need N^i is satisfied after certain actions. So, a solution S^i emerges. Here, S^i forms a contradiction moment by being the negation of Ct^i that can be represented by S^i : *not* Ct^i . So, according to CHAT, the content of S^i will become the content of the goal G^i that will be attached to the problem Ct^i . So, later, when the agent faces Ct^i , the agent will seek to achieve the content of the goal G^i . Again, M^j , representing the content of S^i , will be embodied in the initial components forming S^i . Hence, the problem Ct^i and the sensorimotor content in Cd^j . Here, the interaction of the components does not follow the simple and direct connectionist approach; instead, they are qualitatively synthesized through the negation relationships, leading to the emergence of new entities/structures, i.e., the emerged meanings, representing the abstraction experience, and not merely the quantitative accumulation of the initial components. M^i and M^j belong to a higher and a complex level than the level where the initial components of the contradictions belong to. This is a transition from the tangible to the abstract. For an extensive analysis please see El Maouch and Jin (2022a).

By going back to the twofold transition, the meanings space internalizes the reality components, but also the components, in turn, will conserve in them the outcomes of the internalization, i.e., the meanings (see Fig. 2). So, the sensory-motor content and the needs will embody the emerged meanings. In the case of M^i , they will embody the abstract meaning of “the need for food is not satisfied” (Cr^i), while in the case of M^j , it is the embodiment of the abstract meaning “the need for food is satisfied” (S^i). So, being the content of the perturbation Cr^i , and the solution S^i , the need, the actions, the object of satisfaction, and the sensual content, all will be synthesized, and they became as inseparable in the psychomental plane. They exits through each other. This is a simple demonstration of the inseparability of the experience sides (subjective-objective). There is always a subjective side in the objective, and there is always an objective side in the subjective (see [The Twofold Transition, Extended](#) section).

Thus, the meanings are the realization of the twofold transition, and the internal structure of meanings is the formalization of that transition.

Now, following the emergence of the contradiction-based meanings, as a crucial process in constructing the psychological plane, how such a transition explains the emergence of agency and ownership. Following the above demonstration, it can be said that the reflection process of reality is realized totally through the emergence of meanings.

Agency and Ownership as Abstract Needs-Based Processes

As noted earlier, the needs develop, and new needs emerge. So, the agent does not keep the bio-physiological needs but witnesses the emergence of abstract needs, as well, based on the emergence of subjective components, where the self is at the core (see [Self as a Flow and Objectively Originated](#) and [The Twofold Transition, Extended](#) sections).

Again, we would like to stress that an AI-friendly model of agency and ownership should not be a copy of the one in humans. Therefore, the model should not repeat the same path of development in humans that is investigated in Sect. 3, instead, we should make it AI-friendly. What do we mean by that?

In short, we will employ the endpoint of the development that the psychological plane in humans reaches. It means that we do not need a *self* of the same quality as humans. What we need to model is the needs that appear on the base of *self*. These are the abstract needs that provide the ground for the emergence of the abstract process (in our case the agency). The abstract representation of needs in the system is usually neglected in mainstream studies since the existence of the self is still misunderstood by mainstream directions. What are these abstract needs that will provide the emergence of agency and ownership?

As noted earlier, according to dialectical logic, any system is governed internally by the main laws of contradictions-based development. Again, the internal contradictions are what lead to the qualitative emergence in and of the system, this is from one hand. On another hand, the practical activity is governed by its own structure in terms of being always directed, driven by needs, and motivated by the activity’s

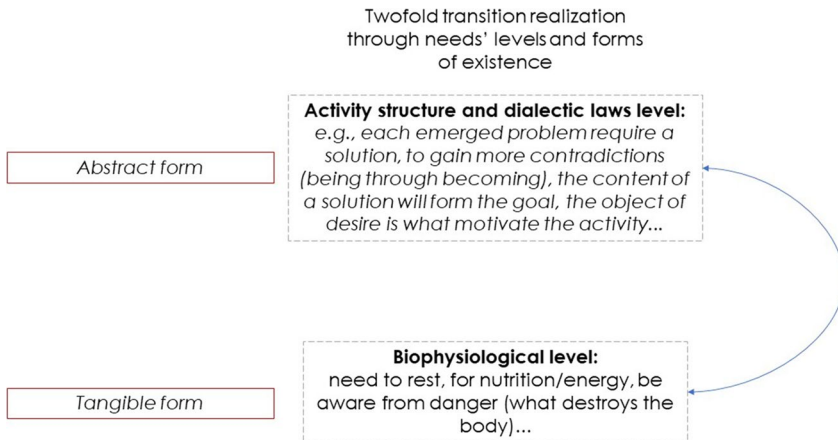


Fig. 3 Twofold transition realization through needs' levels

object, in addition to that what we already introduced regarding that the activity is a process of problems solving as the source of meaning emergence. Therefore, this internal structure of the activity imposes the requirements of the psychological structure, but what is more important is that based on this contradiction-base this structure gain its internal self-movement. Therefore, both the laws of dialectics in addition to the internal structure of the activity (as self-realization entity) became the main source of abstract needs of the system. Next, we will try to formalize some of these needs in order to grasp the emergence of agency as a process.

Thus, abstract needs are the representation of the requirements of the internal structure of the system, exceeding the given bio-physiological organism (e.g., food, rest, and breathing) (see Fig. 3). For instance, being driven by internal contradictions, hence, being, transcendental, the system not only requires that each newly emerged problem needs to be solved but also, the system's development is pushed internally to seek new contradictions. This is the definition of being through becoming (being transcendental) (see [Consciousness Plane, Dynamic System of Meanings \(DSM\), and Contradictions-Based Meanings Emergence](#) section), in contrast with the *autopoietic* one that only considers the system as being *homeostatic*. It is more than “maintaining parameters which are crucial for system's preservation within the tolerable limits” (Klochko, 2008, p. 31) (see El Maouch & Jin, 2022a).

Therefore, by providing the system with this self-expanding aspect, and at the same time, when the system is able to represent it explicitly as a need. The agent's ability to grasp the abstract structure of the activity is crucial for abstract learning and for building a narrative from a subjective point of view (e.g., see El Maouch et al., 2019a, c), because “a robot's narrative allows humans to get an insight into long term human-robot interaction from the robot's perspective” (Moulin-Frier et al., 2017, p. 4).

Thus, two main abstract needs of the system are, **one**, to seek new contradictions, and **two**, each emerging contradiction has to have its own solution. These emerged needs are in turn the source of new contradictions in the system, hence,

new meanings (this is the topic of another interest that is not in our focus at this moment). However, let us focus on how these needs are the crucial source of agency emergence.

Using the language of formalizing the dialectical logic, the first abstract need N^1 that: each problem Ct^x has to have its own solution S , can be represented in an AI-friendly model by **For any** $Ct^x: B \text{ not } A \rightarrow \text{find } S^x: \text{not } (B \text{ not } A)$ that should exist as the desired state DS^1 as the content of the noted need N^1 . The second abstract need N^2 is that new problems should emerge (being through becoming), which can be represented by: increase $Ct(s)$, or: **for** $Ct^x \rightarrow Ct^{x+1}$ (as the desired state DS^2 that is the content of the noted need N^2) (see Table 1). Both N^1 and N^2 represent a formalized AI-friendly version of *intentionality* because the agent, one, is driven by their own internal needs to act, and two, grasping these needs explicitly. We quote:

What we do when we tie action and intentionality together in such a fashion is essentially begin to tell the story of agency. An agency is not just something which just acts. An agent is something that performs actions intentionally (Swanepoel, 2021; p. 85).

Now, in a moment i (in a current state CS^i), when any tangible need N^i is pressing (e.g., food) (as a lack of the desired state DS^i), so a problem Ct^i (the need for food is not satisfied) emerges with the meaning $M^i: CS^i \text{ not } DS^i$. In another moment j (in a current state CS^j), after certain reactions Act^j the body may execute, let us suppose that the emerged need N^i is satisfied and Ct^i is solved (e.g., the caregiver answers the request). The actions that formed the reactions Act^j will gain the meaning of not only satisfying the tangible need N^i (here it is the food) but also gain the meaning of satisfying the abstract need N^1 of the system to find a solution for the emerged problem. Here, not only the exact actions (the operations in the given condition about how many centimeters the agent moved and in which angel) will embody the meaning of the ability to solve any given emerged problem (N^1), but the actions as an abstract concept will gain that meaning as well (when there are *changes* in the parameters of the *motors*, or, being different from 0-Zero which meach that there is *movement*). Why? Here, by going back to that the self is synthesized with the body, and the body is differentiated from the environment, the actions gain the meaning $M^j: \text{me doing an action}$. It is me (self) that achieved this result. By satisfying the

Table 1 Functional content and formalization of agency

| Process | Internal functional content that serves as desired state DS to generate meanings | Formalization |
|---------|--|--|
| Agency | To find a solution (S) for every problem (Ct) | For any $Ct^x: B \text{ not } A \rightarrow \text{find } S^x: \text{not } (B \text{ not } A)$ |
| | Being through becoming: To expand the system by generating new contradictions | Increase $Ct(s)$, or: for $Ct^x \rightarrow Ct^{x+1}$ |

needs, the *self-realization* as a general need is also satisfied because each specific need (tangible or abstract) is the representation of the self-realization need. Self-realization is satisfied through specific needs and specific satisfactions.

Overall, the new meaning M^j is the outcome of satisfying the need N^1 of DS^1 (when each emerged problem requires a solution). Then, M^j is: CS^j not (*not* DS^1). And since the current state CS^j includes the agent's action and the sensory inputs, and DS^1 includes the needs N^1 (as well as the tangible need N^j) so the meaning M^j is the synthesis of these components in a given condition Cr^i . And, M^j can be represented as a solution S^j of Cr^i : not Cr^i ($\neg Cr^i$). Overall M^j is: $(Act^i + Cd^j) \neg (\neg (N^i + N^1 + Cr^i))$ (Eq. 1) (here $(Act^i + Cd^j)$ substitute CS^j , and $(N^i + N^1 + Cr^i)$ substitute DS^1).

Now, regarding the second abstract need N^2 (the system is required to seek new problems), the activity of the body not only leads to needs satisfaction but also leads to the emergence of problems. For instance, simply, the body consumes energy, which means the emergence of new tangible need N^k (e.g., to rest, and nutrition) in moment k . By being so, the activity of the body is itself a source of new problems related to the bio-physiological requirements of the organism. Therefore, due to the body activity, the need N^2 (about increasing the contradictions faced by the system) will be satisfied, and a solution S^k : not Cr^2 , will emerge. Again, the new emerged meaning M^k can be represented by: *not* Cr^2 , or *not* (*not* DS^2). Overall M^k is: $(Act^k + Cd^k) \neg (\neg (N^k + N^2 + Cr^k))$ (Eq. 2) (here $(Act^k + Cd^k)$ substitute CS^k , and $(N^k + N^2 + Cr^k)$ substitute DS^2). In both equations, one and two, we can identify how the bio-physiological needs level is synthesized with the abstract level of needs according to the twofold transition realization through meanings emergence. Here, the agent's desired states have tangible aspects derived from the direct bio-physiological needs, as well as abstract aspects derived from the abstract needs based on the requirements of dialectical development which is the internal content of *being through becoming* and of contradictions-based driven qualitative emergence (N^1 and N^2), as well as derived from the requirements of the activity system regarding how the needs, actions, motives, goals, conditions, sensory input... interact and related in a structured system. The explicit modeling of the noted abstract level through abstract needs (see N^1 and N^2) is crucial to allow the system to grasp its own meta-functions (to our interest, own agency) (see Fig. 3).

As noted earlier, and according to the twofold transition, the initial component of M^j and M^k (e.g., actions and sensory input) will be embodied by M^j and M^k . So, the actions, the needs, and the sensory contents will be perceived later as holders of these meanings. In this way, the experience components are synthesized.

Thus, the meanings space now includes two meanings M^j and M^k representing the active content of the organism regarding the environment based on the satisfaction of the internal abstract requirement of the system as a transcendental one (being through becoming) represented by the abstract needs N^1 and N^2 . Agency here is that the agent grasps the content of being able to change the context to satisfy their own needs (N^1), as well as to expand their own self through facing new problems (N^2), hence, learning new solutions (this is the source of curiosity as well) (see El Maouch & Jin, 2022a). Thus, the agency structure's internal content can be represented with the extended equation: Eq. 1 + Eq. 2, or: $[(Act^i + Cd^j) \neg (\neg (N^i + N^1 + Cr^i))] + [(Act^k + Cd^k) \neg (\neg (N^k + N^2 + Cr^k))]$ (Eq. 3). Equation 3 represents the meaning of being able

to solve problems through active engagement as well as to expand the skills through this engagement by facing new problems/perturbation hence grasping its *autonomy*, as well as *responsible intentionally* about its own actions. This is the formal representation of the system development pushed by its own internal contradictions and seeking *self-realization*.

We consider that this is the realization of the agency being implemented in an AI-friendly model without having an original copy of the individual self. Instead, the internal causal origin of the agency's structure is enough so the agent is equipped with the ability to actively engaged in the environment (mainly by seeking new experiences through increasing the problems (N^2), and by considering that each problem must have a solution (N^1)).

Furthermore, because agency and ownership cannot exist in absolute separation, the noted equations above (Eqs. 1 and 2) should represent the aspect of ownership. Such interaction between ownership and agency is derived from the fact that the activity (executed by the subject) is always directed towards objects. There is no such thing as objectless activity (Leontiev, 1978).

In order that the body (as an active body), that is present through actions, reflect an aspect of ownership should comport two levels. **One** is the differentiation between this body and the environment, bordered by the tangible sensation. So, using the language of dialectical logic formalization, for an AI-friendly model, what is left is to formalize such a statement. Directly, we can say that, in the language of a contradiction-based meaning, the meaning of anything except the self holds the content: “not self”, so the representation of self holds the content: “not “not self” (=self) (self is defined and discovered through other entities). Again, what is the ownership in an AI-friendly model? it is simply a reference so the agent can differentiate between their own self and the environment. What is crucial is that this content will be synthesized in the two equations noted earlier, so such differentiation can support other meanings and other processes. Again, an AI-friendly model is not a copy of human consciousness, so the processes we are targeting in the model are dedicated to achieve certain functions, without being a copy of the one we find in humans. So, regarding ownership, the differentiation between body and environment is enough in the context of artificial systems by attributing the meaning of being or not part of the body.

The **second** level, which is synthesized with the first level (the differentiation between the body and the environment), is that an AI-friendly model should include the embodiment of the meaning of agency in the differentiated part of the body (ownership).

Again, being equipped with artificial tactile sensors will increase the system's interaction with the environment/objects, hence, more aspects of the emerged meanings we have then; however, what is required as well is to have certain marks on the surface of the body. These marks will be part of the sensory content included in the equations of the agency in addition to the tactile input. These inputs will be the source of, one, differentiating the body borders, and two, the objects' dimensions and scales. However, following the contradictions-based meanings, in order that these inputs (from a given condition) originate a meaning of ownership, they have to be evaluated by the functional content M^{own} : “not body” (or “not ‘not

body”’, in the case of refereeing to body). Here, the *body* can be simply a given symbol, e.g., *Bd* (what is important here is not if the agent feels a real body, as long as the given designed system is realizing the required functions). The contradiction between *belonging* and *not belonging* is usually noted in literature without formalizing it into the concepts of dialectics, instead, it stays on the formal linguistic description such as in the following statement: “we can divide sensory experiences into the ones which do relate to the self and the ones which do not”, so “self-as-object...provides a guiding principle for distinguishing between self and non-self” (Woźniak, 2018, p. 4). However, in our proposed model, both agency and ownership are molded together in one structure while one appears through the other. The tactile and the surface markers content will be embodied by the meaning “not ‘not body’” (being part of the body), and other surfaces will embody the meaning “not body” (not belonging to the body). The meanings generated here will be included in the equation of agency noted earlier. In this way, the ownership is generated via another path (another functional content) that of agency, but both are synthesized at the end from the point of view of practical actualization/realization.

Moreover, being synthesized in the agency equation, ownership is not merely sensual *passive* input, rather, it represents the active content of an active body. Why? Again, the initial content of the emerged meaning will embody the generated meaning. So, being part of the equation, the ownership meaning as well will embody the meaning of agency that is generated through the agency equation. It is about owning an “active body” synthesized with needs (a needy body) (see Fig. 4). Therefore, the extension of Eq. 3 (by including the ownership meaning M^{own}) will take the following structure: $[(Act^i + Cd^i) + M^{own}] \neg (\neg (N^i + N^1 + Ct^i)) + [(Act^k + Cd^k) + M^{own}] \neg (\neg (N^k + N^2 + Ct^k))$ (Eq. 4). As noted earlier, the ownership meaning (being or not being part of the body) forms an abstract input to the agency equation. The inclusion of the evaluative component (meaning) M^{own} in the agency meaning, that is the synthesis of M^j and M^k , is a representation of the concept of *functional blocks*. M^{own} is to a certain extent an example of a participating component in the meanings’ block included in the self-regulation model that is introduced in the Systematic Structural Activity Theory (SSAT).

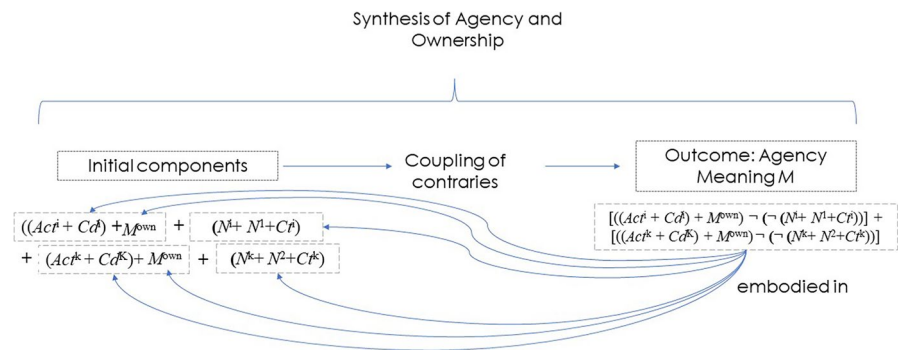


Fig. 4 Twofold transition synthesis of agency and ownership

Just to note that SSAT is an application of CHAT in ergonomics and Human-Computer Interaction (HCI) (e.g., see Bedny & Karwowski, 2006; Bedny et al., 2014). The meanings' blocks represent the complexity of how the dynamic system of meanings (DSM) governs the perceptive processes (see [Consciousness Plane, Dynamic System of Meanings \(DSM\), and Contradictions-Based Meanings Emergence](#) section), where each meaning is not and cannot be separated from other meanings in the system. For instance, our proposition about the agency's internal structure reflects the synthesis of needs, desires, and functional actions (since these actions are not merely physical ones, but achieve certain functions related to goals that are grasped by the agent) that is for some researcher as accounts for the agency in an AI-friendly model (e.g., see Swanepoel, 2021). Again, what differs in the proposed model is that the tactile inputs are synthesized with other experience components (actions, needs, other sensory inputs, etc.) via the crucial negation relationship *not* (\neg), first, in the ownership meaning M^{own} , and second, in the agency equation $[(\text{Act}^i + \text{Cd}^i) + M^{\text{own}}] \rightarrow (\neg (N^i + N^1 + \text{Cr}^i))$ and $[(\text{Act}^k + \text{Cd}^k) + M^{\text{own}}] \rightarrow (\neg (N^k + N^2 + \text{Cr}^k))$. Thus, it is the negation relationship that forms the core regarding how components are qualitatively synthesized to generate abstract and functional meanings (see [Consciousness Plane, Dynamic System of Meanings \(DSM\), and Contradictions-Based Meanings Emergence](#) section). Furthermore, the equation could include more meanings blocks (other meanings related to other perceptive functions that may be designed and implemented, such as grasping, directions, etc.) that form the internal structure of the condition Cd (the sensory inputs); however, for the goal of simplicity, we shall only focus on the ownership meaning block.

We are not setting the final version of these processes, but refereeing to their general structure and internal functional contents, as well as their interaction in a unified framework. Of course, one can always extend and increase their complexity by adding more meanings related to other mental processes where both agency and ownership are included and represented as part of a specific (reach, grasp, navigate...). However, our focus in this discussion is the interaction between agency and ownership, as well as their internal content expressed as contradictions-based processes. Overall, what is important is a formalized general version of how the agency and ownership are internally related is introduced.

We can see that since we allowed the system to be based on the meanings' space, the illusions that are in the phenomenal experience in the case of humans can be neglected here (e.g., see Aron et al., 1992; Braun et al., 2018; Woźniak, 2018), because, as noted earlier (see [“I”: the Subjective Side of the Subjective Existence and Twofold transition, Agency and SoA, Ownership and SoO](#)) the confusion emerges due to the interrelation of reflected reality in the space of self and “I”, resulting that the objects in the objective environment (and the mental plane) are felt as injected with the sense of self and “I”. But since in the proposed AI-friendly model we already excluded the gap between the space of meanings and the space of senses (see [The Twofold Transition, Extended](#) section), the model is close to the objective meanings, only when the internal content of these meanings is grasped, hence, preventing illusions from being generated. For instance, although the agency and the body are molded to form the agentive body, still, the body is not mixed phenomenologically with the environment. Such a distancing of the body from being mixed with the environment does not mean that the objects of the environment are totally

de-subjectified, but as we noted in “[The Twofold Transition, Extended](#)” section, what is embodied in the objects is the agentive aspect of the body (an object that *I* once acted on/with). So, the concept of objects as extensions of the body is not the same as the phenomenological illusion. In this case, the agent/system can still differentiate these objects as not part of its own, but at the same time perceive the agency in them.

The previous does not exhaust what such a model of formalization can provide. For instance, as we saw earlier, we only referred to the physical type of actions conducted by the agent/system, but it is necessary to say that mental/cognitive actions (and tools) should be included if the design of the system requires such an explicit grasp of these mental actions. However, for the sake of limited space, we shall let go of this part for another occasion when a discussion of mental actions and their role in activity is required, by keeping our focus here on formalizing agency and ownership.

Some Implications

Moreover, the proposed model, as being equipped with built-in *needs* (see Formalization of Agency and Ownership section) to be fulfilled and being “designed” by humans, represents to a certain level the philosophical argument of Popa (2021) that the agency in an artificial system is the extension of human goals, hence, the extension of human agency, which, in turn, contribute to the ongoing ethical debate in the field regarding the role of agency in responsibility from the standpoint of the mutual and complex understanding of ethics and responsibility which goes beyond the narrow individualistic agent-based concept of responsibility toward a wider sociohistorical, contextual, and systematic view, without neglecting that revealing the internal structure of the agency and its design increase the transparency of the system which is considered as an important aspect in ethical debate in AI (e.g., see Andrada et al., 2022). Also, our investigation reveals that the agency as we experience it in humans is still out of reach considered the reach in AI taking into consideration not only the *powerful* bio-neurophysiological infrastructure (the brain) where our consciousness is realized that goes back to several millions of years of development but also considering the richness and life-long aspects of human experience. The previous *ethical* implication is one of the various *practical* implications of how the agency and ownership contribute to the field of AI including human-AI interaction (especially in terms of collaboration, personification, and socialization) (e.g., see Sundar, 2020), providing the crucial ground for active learning and curiosity (some aspect of being autonomy)(Poltronieri et al., 2020), as well as modeling other agents by being able to grasp and discover the decision factors, the efficient computational implementation, and being able to abstract the behavior by grasping its internal content with distance from the tangible conditions of the experience, taking into consideration other agents’ goals, belief, and actions especially since one can already model own autonomy and goes beyond the formal and sensual experience (e.g., see Albrecht & Stone, 2018).

Moreover, the participation of the ownership process in the internal structure of the *agency* is not a single case in the psychological plane. Instead, all the mental processes are interrelated and synthesized following the unified aspect of the

reflection and the internal dynamic of the system (see El Maouch & Jin, 2022a). In the emergence of meanings, we notice the realization of Leontiev (1978)'s argument that the activity is a molar. However, our focus here was on both agency and ownership, although such isolation among mental processes does not exist in reality. So, in addition to other works (e.g., El Maouch et al., 2019a, b, c; El Maouch & Jin, 2022a), this paper tries to introduce not only the unified internal structure of the mind, which is based on the dynamic system of meanings through contradictions coupling, that is the qualitative path to overcome the quantitative tendencies in the field, hence, bridge the gap that is the source of the main challenges in the field (e.g., abstraction, and generalization), but also introduce how one cannot isolate one mental process from the other, which is a formalized clue about the unified aspect of the mind.

Conclusion

This is not an exhaustive attempt to look into the implications of the analysis of self and “I” on AI from the position of CHAT and dialectic. Instead, it is an introduction to a qualitative point of view apart from mainstream tendencies in the field, on one hand, the quantitative and connectionist, and the phenomenological ones on the other. Also, this paper aims to establish a preliminary ground to formalize psychological processes through the investigation of the objective activity where these processes originated and are anchored, and it is where the functional ground of the mental. Our work is a formalized extension of the work done by CHAT researchers, and other formalization attempts of dialectic by contextualizing such formalization in the field of AI, and especially robotics, taking into consideration crucial elements of systems and agents such as needs, goals, actions, etc. (e.g., El Maouch et al., 2019a, b, c; El Maouch & Jin, 2022a).

Furthermore, by being the outcome of a twofold transition, we tried to uncover the causal origin of agency and ownership, as well as their interrelation as both are synthesized through meanings emergence, where ownership is a specific aspect of how the agency is realized. It is the agency of someone (self in the body) regarding something (objects) (see Activity, Mind, Dynamic System of Meanings (DSM), and Self section). We consider that the introduced methodological and formalized position is the realization of the arguments that the internal content of the agency is about initiating and controlling its own actions when affecting the external world that is deeply interrelated with the ownership of the agent's body (or body parts), thoughts, and feelings. Moreover, we consider that in this paper the argument about SoA and SoO, which represents the grounding of the sense of self is also realized by revealing that the production of agency and ownership follow the space of self (meanings), while the sense of agency (SoA) and ownership (SoO) is related to the space of “I” (senses-making), hence, for AI systems, both spaces can be considered as one since the design of the system allows it to catch the formed meanings based on their internal structure and not based on the *illusive* phenomenological level that leads to making senses of it which may be distanced

from these meanings (e.g., see Stetsenko & Arievitch, 2004; Moore, 2016; Braun et al., 2018; Gozli, 2019; Legaspi et al., 2019; Stetsenko, 2019).

Overall, the discussion of the implication was somehow straightforward and does not exhaust the richness of the twofold transition (we consider this modeling of various spaces of activity supports the functional analysis by Oliver Sacks regarding his investigation about the functional interaction between the body and mental processes; e.g., see Stone et al., 2012), which we will expand it in upcoming works.

Finally, this paper presents several contributions we consider deserve to mention again. One is the formal representation of the three spaces of existence and the interrelation which exceed the two levels' representation that is usually promoted (see Fig. 1). The second is the formalization of the twofold transition using the internal structure of contradictions-based meaning (see Fig. 2). The third is the attempt to *inject life* in how mental spaces and processes are formed by revealing their causal and contextual origin. The fourth is to propose a formal model of agency (and ownership) that can be implemented in artificial systems in order to handle the current challenges in the field, especially the matter of autonomy and interaction from a subjective standpoint, hence, providing the ground to eliminate the confusion and bottleneck situation derived from the quantitative, mechanistic, and connectionist mainstream tendencies, that are usually either on the pure biologist and sensual pole or on the phenomenological pole. Overall, this paper continues a line of analysis in previous work about fortifying the position of a qualitative methodology in the formalization of intelligence's base, by taking into consideration that the cornerstone in such an attempt is the contradiction-based dynamic system of meanings which allows us to bridge the epistemological and ontological gap and rupture between mind-body, abstract-tangible, subject-object, etc. away from the mainstream mechanistic and quantitative tendencies. Still, this is only a small step in discovering how such a qualitative position may shift not only the practical level of AI but also the ethical and sociological level in understanding the limits of the effects of AI, by including not only human collaborators in it, but also the whole sociohistorical structure when the agency is considered a sociohistorical process and not as an absolute entity in its own that may appear from the *dark side of the void*.

Author Contributions M.E. was principally responsible for the text in terms of epistemological, theoretical, and methodological proposals as well as developing the alternative guidelines (including figures and tables). Z.J. contributed to the review and the historical investigation and provided methodological and literature suggestions. Both authors contributed to the manuscript's final form.

Funding This study was partially supported by Program for Science and Technology Development of Henan Province (222102310686) and Talents Program of the Ministry of Science and Technology of the PRC.

Data Availability Data sharing not applicable to this article as no datasets were generated or analyzed during the current study.

Declarations

Competing Interests The authors declare no competing interests.

References

- Albrecht, S. V., & Stone, P. (2018). Autonomous agents modelling other agents: A comprehensive survey and open problems. *Artificial Intelligence*, 258, 66–95. <https://doi.org/10.1016/j.artint.2018.01.002>
- Albrechtsen, H., Andersen, H. H., Bødker, S., & Pejtersen, A. M., A. M (2001). *Affordances in activity theory and cognitive systems engineering*. Risø National Laboratory. Retrieved from <https://core.ac.uk/download/pdf/13771112.pdf>. Accessed 1 Sept 2022.
- Allé, M. C., Potheegadoo, J., Köber, C., Schneider, P., Coute, R., Habermas, T., ... Bern, F. (2015). Impaired coherence of life narratives of patients with schizophrenia. *Scientific Reports*, 5, 12934. <https://doi.org/10.1038/srep12934>
- Andrada, G., Clowes, R. W., & Smart, P. R. (2022). Varieties of transparency: Exploring agency within AI systems. *AI & Society*, 34, 1213–1242. <https://doi.org/10.1007/s00146-021-01326-6>
- Aron, A., Aron, E. N., & Smollan, D. (1992). Inclusion of other in the self scale and the structure of interpersonal closeness. *Journal of Personality and Social Psychology*, 63(4), 596–612. <https://doi.org/10.1037/0022-3514.63.4.596>
- Babaeva, Y. D., Berezanskaya, N. B., Kornilova, T., Vasilyev, I. A., & Voiskounsky, A. E. (2013). Contribution of Oleg K. Tikhomirov to the methodology, theory and experimental practice of psychology. *Psychology in Russia: State of the Art*, 6(4), 4–23. <https://doi.org/10.11621/pir.2013.0401>
- Bedny, G., & Karwowski, W. (2006). *A systemic-structural theory of activity: Applications to human performance and work design*. CRC Press. <https://doi.org/10.1201/9781420009743>
- Bedny, G. Z., Karwowski, W., & Bedny, I. (2014). *Applying systemic-structural activity theory to design of human-computer interaction systems* (1st ed.). CRC Press. <https://doi.org/10.1201/b17975>
- Benke, C., Autenrieth, L. K., Asselmann, E., & Pané-Farré, C. A. (2020). Lockdown, quarantine measures, and social distancing: Associations with depression, anxiety and distress at the beginning of the COVID-19 pandemic among adults from Germany. *Psychiatry Research*, 293, 113462. <https://doi.org/10.1016/j.psychres.2020.113462>
- Bickhard, M. H., & Terveen, L. (1996). *Foundational issues in artificial intelligence and cognitive science: Impasse and solution* (109 vol.). Elsevier.
- Bolis, D., & Schilbach, L. (2020). I interact therefore I am’: The self as a historical product of Dialectical Attunement. *Topoi*, 39, 521–534. <https://doi.org/10.1007/s11245-018-9574-0>
- Braun, N., Debener, S., Spychala, N., Bongartz, E., Sörös, P., Müller, H. H., & Philippen, A. (2018). The senses of agency and ownership: A review. *Frontiers in Psychology*, 9(535), 1–17. <https://doi.org/10.3389/fpsyg.2018.00535>
- Carter, M. (2007). *Minds and computers: An introduction to the philosophy of artificial intelligence*. Edinburgh University Press.
- Charles, D. (2021). *The undivided self: Aristotle and the 'mind-body problem'*. Oxford University Press.
- Chew, Q. H., Wei, K. C., Vasoo, S., Chua, H. V., & Sim, K. (2020). Narrative synthesis of psychological and coping responses towards emerging infectious disease outbreaks in the general population: Practical considerations for the COVID-19 pandemic. *Singapore Medical Journal*, 61(7), 350. <https://doi.org/10.11622/smedj.2020046>
- Counet, J. M. (2012). La formalisation de la dialectique de Hegel. Bilan de quelques tentatives. *The Log Analyst*, 55, 205–227.
- Cruse, H., Dean, J., & Ritter, H. (Eds.). (2000). *Prerational Intelligence: Adaptive behavior and Intelligent Systems without Symbols and Logic* (1 vol.). Springer Science & Business Media. <https://doi.org/10.1007/978-94-010-0870-9>
- Dafermos, M. (2014). Vygotsky’s analysis of the crisis in psychology: Diagnosis, treatment, and relevance. *Theory & Psychology*, 24(2), 147–165. <https://doi.org/10.1177/0959354314523694>
- Davidson, L. (1993). Story telling and schizophrenia: Using narrative structure in phenomenological research. *The Humanistic Psychologist*, 21(2), 200–220. <https://doi.org/10.1080/08873267.1993.9976919>
- Davydov, V. V. (1990). *Types of generalization in instruction: Logical and psychological problems in the structuring of School Curricula* (Vol. 2). National Council of Teachers of Mathematics. (J. Teller, Trans.).
- Dennett, D. C. (2017). *Brainstorms: Philosophical essays on mind and psychology*. MIT Press. <https://doi.org/10.7551/mitpress/11146.001.0001>
- Drescher, G. L. (1986). Genetic AI: Translating piaget into LISP. *Instructional Science*, 14, 357–380. <https://doi.org/10.1007/bf00051828>

- Drescher, G. L. (2003). *Made-up minds: a constructivist approach to artificial intelligence*. MIT press.
- Dreyfus, H. L. (1981). From micro-worlds to knowledge representation: AI at an impasse. In J. Haugel (Ed.), *Min Design* (pp. 161–204). MIT Press.
- Ekbia, H. R. (2008). *Artificial dreams: The quest for non-biological intelligence*. Cambridge University Press. <https://doi.org/10.1017/CBO9780511802126>
- El Maouch, M., & Jin, Z. (2022a). Artificial Intelligence inheriting the historical crisis in psychology: An epistemological and methodological investigation of challenges and alternatives. *Frontiers in Psychology*, *13*(781730), 1–20. <https://doi.org/10.3389/fpsyg.2022.781730>
- El Maouch, M., & Jin, Z. (2022b). Activity system, schizotypal personality, and mentalization: A study between halted activity and COVID-19 conducted in Henan, China. *Frontiers in Public Health*, *10*(930842), 1–21. <https://doi.org/10.3389/fpubh.2022.930842>
- El Maouch, M., Wang, S., Xiaosong, G., Lu, T., Cai, Y., & Saab, M. (2019a). Activity theory based model for robotics tasks learning and functional reflection. *Proceedings of the 31st Chinese Control and Decision Conference* (pp. 3572–3579). IEEE. <https://doi.org/10.1109/CCDC.2019.8832472>
- El Maouch, M., Wang, S., & Xiaosong, G. (2019b). Contradictions based qualitative development in robotics, for meanings emergence and functional reflection. *Proceedings of the 14th China Intelligent Robots Conference (CIRC)*, (pp. 18–24). Chongqing.
- El Maouch, M., Wang, S., & Xiaosong, G. (2019c). Contradictions-based development in robotics: meanings emergence and concepts formation for communication, and words acquisition. *Chinese Automation Congress, CAC* (pp. 1990–1995). IEEE. <https://doi.org/10.1109/CAC48633.2019.8996948>
- Feuerbach, L. (1989). *The essence of christianity* (G. Eliot, Trans.). Prometheus Books.
- Frankish, K. (2021). Panpsychism and the depsychologization of consciousness. *Aristotelian Society Supplementary Volume*, *95*(1), 51–70. <https://doi.org/10.1093/arisup/akab012>
- Froese, T. (2007). On the role of AI in the ongoing paradigm shift within the cognitive sciences. In M. Lungarella, F. Iida, J. Bongard, & R. Pfeifer (Eds.), *50 years of Artificial Intelligence* (Vol. 4850, pp. 63–75). Springer. https://doi.org/10.1007/978-3-540-77296-5_7
- Girãoa, P. S., Ramos, P., Postolache, P. M., & Pereira, J. M. (2013). Tactile sensors for robotic applications. *Measurement*, *46*(3), 1257–1271. <https://doi.org/10.1016/j.measurement.2012.11.015>
- Gjorgjioska, M. A., & Tomicic, A. (2019). The crisis in social psychology under neoliberalism: Reflections from social representations theory. *Journal of Social Issues*, *75*(1), 169–188. <https://doi.org/10.1111/josi.12315>
- Goertzen, J. (2008). On the possibility of unification: The reality and nature of the crisis in psychology. *Theory & Psychology*, *18*, 829–852. <https://doi.org/10.1177/0959354308097260>
- Goff, P. (2009). Why Panpsychism doesn't help us explain consciousness. *Dialectica*, *63*(3), 289–311. <https://doi.org/10.1111/j.1746-8361.2009.01196.x>
- Gopal, A., Sharma, A. J., & Subramanyam, M. A. (2020). Dynamics of psychological responses to COVID-19 in India: A longitudinal study. *PLoS One*, *15*(10), e0240650. <https://doi.org/10.1371/journal.pone.0240650>
- Gozli, D. (2019). Sense of Agency. *Experimental psychology and human agency* (pp. 137–164). Springer. https://doi.org/10.1007/978-3-030-20422-8_7
- Group of contributors. (2020, Sept 12). *Philosophy of self*. Retrieved from Psychology WIKI: https://psychology.fandom.com/wiki/Philosophy_of_self. Accessed 14 Sept 2022.
- Gribanov, D. (1981). *Materialism and Dialectic in Einstein Thought*. Al Farabi.
- Guessoum, S. B., Lachal, J., Radjack, R., Carretie, E., Minassian, S., Benoit, L., & Moro, M. R. (2020). Adolescent psychiatric disorders during the COVID-19 pandemic and lockdown. *Psychiatry research*, *113264*. <https://doi.org/10.1016/j.psychres.2020.113264>
- Hagell, A., Peck, S. C., Zarrett, N., Giménez-Nadal, J. I., & Symonds, J. (2012). Trends in adolescent time-use in the United Kingdom. *Changing adolescence: Social trends and mental health*, 47–73. <https://doi.org/10.2307/j.ctt9qgs8z>
- Haselager, W. F. (2005). Robotics, philosophy and the problems of autonomy. *Pragmatics & Cognition*, *13*(3), 515–532. <https://doi.org/10.1075/pc.13.3.07has>
- Haselager, W. F., & Gonzalez, M. E. (2007). Mechanicism and autonomy: What can robotics teach us about human cognition and action? *Pragmatics & Cognition*, *15*(3), 407–412. <https://doi.org/10.1075/pc.15.3.02has>
- Haugeland, J. (ed.) (1997). *Mind design II: Philosophy, psychology, artificial intelligence*. MIT Press.
- Hodgekins, J., French, P., Birchwood, M., Mugford, M., Christopher, R., Marshall, M., & Fowler, D. (2015). Comparing time use in individuals at different stages of psychosis and a non-clinical

- comparison group. *Schizophrenia Research*, 161(2–3), 188–193. <https://doi.org/10.1016/j.schres.2014.12.011>
- Huynh, T. V., Bekrater-Bodmann, R., Fröhner, J., Vogt, J., & Beckerle, P. (2019). Robotic hand illusion with tactile feedback: Unravelling the relative contribution of visuotactile and visuomotor input to the representation of body parts in space. *PLoS One*, 14(1), e0210058. <https://doi.org/10.1371/journal.pone.0210058>
- Hyman, L. (2012). Vygotsky's crisis: Argument, context, relevance. *Studies in History and Philosophy of Biological and Biomedical Sciences*, 43, 473–482. <https://doi.org/10.1016/j.shpsc.2011.11.007>
- James, W. (1890). *The principles of psychology*. Holt and Company.
- Kappasov, Z., Corrales, J. A., & Véronique, P. (2015). Tactile sensors for robotic applications. *Robotics and Autonomous Systems*, 74(Part A), 195–220. <https://doi.org/10.1016/j.robot.2015.07.015>
- Kelley, T. D., & Cassenti, N. D. (2011). Theoretical explorations of cognitive robotics using developmental psychology. *New Ideas in Psychology*, 29, 228–234. <https://doi.org/10.1016/j.newideapsych.2009.07.002>
- Kiltner, K., & Ehrsson, H. (2017). Body ownership determines the attenuation of self-generated tactile sensations. *Proceedings of the National Academy of Sciences of the United States of America*, 114(31), 8426–8431. <https://doi.org/10.1073/pnas.1703347114>
- Klochko, V. Y. (2008). Modern psychology: systems meaning of a paradigm shift. *Psychology in Russia: State of the Art*, 1, 25–35. <https://doi.org/10.11621/pir.2008.0002>
- Kober, J., Bagnell, J. A., & Peters, J. (2013). Reinforcement learning in robotics: a survey. *International Journal of Robotics Research*, 32, 1238–1274. <https://doi.org/10.1177/0278364913495721>
- Kosok, M. (1966). The formalization of Hegel's dialectical logic: Its formal structure, logical interpretation and intuitive foundation. *International Philosophical Quarterly*, 6, 596–631. <https://doi.org/10.5840/ipq19666456>
- Kosok, M. (1976). The systematization of dialectical logic for the study of development and change. *Human Development*, 19(6), 325–350. <https://doi.org/10.1159/000271540>
- Lakoff, G., & Johnson, M. (1999). *Philosophy in the flesh: the embodied mind and its challenge to western thought*. Basic Books.
- Legaspi, R., Zhengqi, H., & Taro, T. (2019). Synthetic agency: Sense of agency in artificial intelligence. *Current Opinion in Behavioral Sciences*, 29, 84–90. <https://doi.org/10.1016/j.cobeha.2019.04.004>
- Leontiev, A. N. (1978). *Activity, Consciousness, and Personality*. Prentice-Hall. Retrieved from <https://www.marxists.org/archive/leontev/works/activity-consciousness.pdf>. Accessed 5 Aug 2022.
- Leontyev, A. (1981). *Problems of the development of the mind*. Progress.
- Luria, A. R. (1966). *Higher cortical functions in man*. Springer. <https://doi.org/10.1007/978-1-4684-7741-2>
- Luria, A. R. (1976). *Cognitive development: Its cultural and social foundations*. Harvard University Press.
- Lysaker, P. H., Ringer, J., Maxwell, C., McGuire, A., & Lecomte, T. (2010). Personal narratives and recovery from schizophrenia. *Schizophrenia Research*, 121(1–3), 271–276. <https://doi.org/10.1016/j.schres.2010.03.003>
- Ma, Z., Zhao, J., Li, Y., Chen, D., Wang, T., Zhang, Z., & Liu, X. (2020). Mental health problems and correlates among 746 217 college students during the coronavirus disease 2019 outbreak in China. *Epidemiology and Psychiatric Sciences*, 29(e181), 1–10. <https://doi.org/10.1017/S2045796020000931>
- Marx, K., & Engels, F. (1996). *Marx and Engels collected works vol 35: Karl Marx Capital: Volume 1*. Lawrence and Wishart.
- Marx, K., & Engels, F. (1997). *Marx and Engels collected works vol 36: Karl Marx capital: Volume 2*. Lawrence and Wishart.
- Marx, K., & Engels, F. (1998). *Marx & Engels collected works vol 37: Karl Marx capital: Volume 3*. Lawrence & Wishart.
- Massad, I., Al-Taher, R., Massad, F., Al-Sabbagh, M. Q., Haddad, M., & Abufaraj, M. (2020). The impact of the COVID-19 pandemic on mental health: Early quarantine-related anxiety and its correlates among Jordanians. *Eastern Mediterranean Health Journal*, 26(10), 1165–1172. <https://doi.org/10.26719/emhj.20.115>
- Medina, J., & Coslett, H. B. (2010). From maps to form to space: Touch and the body schema. *Neuropsychologia*, 48, 645–654. <https://doi.org/10.1016/j.neuropsychologia.2009.08.017>
- Meijers, J., Harte, J. M., Jonker, F. A., & Meynen, G. (2015). Prison brain? Executive dysfunction in prisoners. *Frontiers in Psychology*, 6, 43. <https://doi.org/10.3389/fpsyg.2015.00043>

- Menant, C. (2003). Information and meaning. *Entropy*, 5(2), 193–204. Retrieved from <https://www.mdpi.org/entropy/papers/e5020193.pdf>. Accessed 10 Jun 2021.
- Menant, C. (2011). Computation on information, meaning and representations. An evolutionary approach. In D. Gordana (Ed.), *Information and computation. Essays on Scientific and Philosophical understanding of foundations of information and computation* (pp. 255–286). World Scientific Publishing Co. Pte. Ltd. https://doi.org/10.1142/9789814295482_0010
- Menant, C. (2015). *Life, local constraints and meaning generation. An evolutionary approach to cognition*. Retrieved from <https://philpapers.org/rec/MENLLC-3>. Accessed 10 Jun 2021.
- Menant, C. (2020). *Introduction to a systemic theory of meaning*. Retrieved from <https://crmenant.free.fr/MGS-Short.Paper-March.2020>. Accessed 11 Jun 2021.
- Mikhailov, F. T., & Daghli, R. (1980). *The riddle of the self*. Progress Publishers.
- Miller, A. V. (Ed.). (1977). *Phenomenology of spirit*. Oxford University Press.
- Moore, J. W. (2016). What is the sense of agency and why does it matter? *Frontiers in Psychology*, 7(1272), 1–9. <https://doi.org/10.3389/fpsyg.2016.01272>
- Moulin-Frier, C., Fischer, T., Petit, M., Pointeau, G., Puigbo, J. Y., Pattacini, U., & Verschure, P. F. (2017). DAC-h3: A proactive robot cognitive architecture to acquire and express knowledge about the world and the self. *IEEE Transactions on Cognitive and Developmental Systems*, 10(4), 1005–1022. <https://doi.org/10.1109/TCDS.2017.2754143>
- Munsterberg, H., Ribot, T., Janet, P., Jastrow, J., Hart, B., & Prince, M. (1910). *Subconscious Phenomena*. Rebman, Ltd. Retrieved from <https://www.nevillegoddardbooks.com/PDF3/Subconscious%20Phenomena%20-%20Munsterberg,%20Ribot,%20Janet,%20Jastrow,%20Hart,%20Prince%201910.pdf>. Accessed 15 Sept 2022.
- Odrizola-González, P., Planchuelo-Gómez, Á, Irurtia, M. J., & de Luis-García, R. (2020). Psychological effects of the COVID-19 outbreak and lockdown among students and workers of a spanish university. *Psychiatry Research*, 290, 113108. <https://doi.org/10.1016/j.psychres.2020.113108>
- Pandey, D., Bansal, S., Goyal, S., Garg, A., Sethi, N., Pothiyill, D. I., & Sethi, R. (2020). Psychological impact of mass quarantine on population during pandemics—the COVID-19 lock-down (COLD) study. *PLoS One*, 15(10), e0240501. <https://doi.org/10.1371/journal.pone.0240501>
- Poltronieri, F. A., DiBlasi, J., Castellanos, C., Kang, E., & Smith, L. (2020). Agency & autonomy: Intersections of artificial intelligence and creative practice. *ISEA2020: 26th International Symposium on Electronic Art*.
- Popa, E. (2021). Human goals are constitutive of agency in Artificial Intelligence (AI). *Philosophy & Technology*, 34, 1731–1750. <https://doi.org/10.1007/s13347-021-00483-2>
- Powers, T. A., Koestner, R., & Zuroff, D. C. (2007). Self-criticism, goal motivation, and goal progress. *Journal of Social and Clinical Psychology*, 26(7), 826–840. <https://doi.org/10.1521/jscp.2007.26.7.826>
- Proctor, H. (2022). Astronomers of the inward: On the histories and case histories of Alexander Luria and Oliver Sacks. *Studies in East European Thought*, 74, 39–55. <https://doi.org/10.1007/s11212-021-09418-1>
- Quintino-Aires, J. (2016). Emptiness in psychological science and practice. *Psychology in Russia*, 9(4), 53. <https://doi.org/10.11621/pir.2016.0402>
- Riegel, K. (1979). *Foundations of dialectical psychology*. Academic. <https://doi.org/10.1016/C2013-0-11367-0>
- Roe, D., & Davidson, L. (2005). Self and narrative in schizophrenia: Time to author a new story. *Medical Humanities*, 31(2), 89–94. <https://doi.org/10.1136/jmh.2005.000214>
- Roth, W. M. (2014). Reading activity, consciousness, personality dialectically: Cultural-historical activity theory and the centrality of society. *Mind Culture and Activity*, 21(1), 4–20. <https://doi.org/10.1080/10749039.2013.771368>
- Schütz-Bosbach, S., Musil, J. J., & Haggard, P. (2009). Touchant-touché: The role of self-touch in the representation. *Consciousness and Cognition*, 18, 2–11. <https://doi.org/10.1016/j.concog.2008.08.003>
- Semenova, N. (2020). Psychosocial treatments for schizophrenia: Insights from russian theoretical psychology. *General Psychiatry*, 33(5), 1–4. <https://doi.org/10.1136/gpsych-2020-100234>
- Serino, A., & Haggard, P. (2010). Touch and the body. *Neuroscience and Biobehavioral Reviews*, 34, 224–236. <https://doi.org/10.1016/j.neubiorev.2009.04.004>
- Shi, L., Lu, Z. A., Que, J. Y., Huang, X. L., Liu, L., Ran, M. S., & Bao, Y. P. (2020). Prevalence of and risk factors associated with mental health symptoms among the general population in China during the Coronavirus Disease 2019 pandemic. *JAMA Network Open*, 3(7), e2014053. <https://doi.org/10.1001/jamanetworkopen.2020.14053>

- Singh, S., Roy, D., Sinha, K., Parveen, S., Sharma, G., & Joshi, G. (2020). Impact of COVID-19 and lockdown on mental health of children and adolescents: A narrative review with recommendations. *Psychiatry Research*, (293), 113429. <https://doi.org/10.1016/j.psychres.2020.113429>
- Smith, J. A., Harré, R., & Van Langenhove, L. (Eds.). (1995). *Rethinking psychology* (3 vol.). Sage.
- Stetsenko, A. (2019). Radical-transformative agency: Continuities and contrasts with relational agency and implications for education. *Frontiers in Education*, 4(148), 1–13. <https://doi.org/10.3389/educ.2019.00148>
- Stetsenko, A., & Arievitch, I. M. (2004). The self in cultural-historical activity theory: Reclaiming the unity of social and individual dimensions of human development. *Theory & Psychology*, 14(4), 475–503. <https://doi.org/10.1177/0959354304044921>
- Stojanov, G. (2009). History of usage of Piaget's theory of cognitive development in AI and robotics: a look backwards for a step forwards. In: *Proceedings of the Ninth International Conference on Epigenetic Robotics*, (pp 243–244). Lund University Cognitive Science.
- Stone, J., Perthen, J., & Carson, A. J. (2012). A Leg to stand on' by Oliver Sacks: A unique autobiographical account of functional paralysis. *Journal of Neurology Neurosurgery & Psychiatry*, 83, 864–867. <https://doi.org/10.1136/jnnp-2012-302800>
- Sundar, S. S. (2020). Rise of machine agency: A framework for studying the psychology of human–AI interaction (HAI). *Journal of Computer-Mediated Communication*, 25(1), 74–88.
- Swanepoel, D. (2021). Does artificial intelligence have agency? In R. W. Clowes, K. Gärtner, & I. Hipólito (Eds.), *The mind-technology problem, studies in brain and mind* (Vol. 18, pp. 83–104). Springer. https://doi.org/10.1007/978-3-030-72644-7_4
- Tagini, A., & Raffone, A. (2010). The 'I' and the 'me' in self-referential awareness: A neurocognitive hypothesis. *Cognitive Processing*, 11, 9–20. <https://doi.org/10.1007/s10339-009-0336-1>
- Tee, M. L., Tee, C. A., Anlcan, J. P., Aligam, K. J., Reyes, P. W., Kuruchittham, V., & Ho, R. C. (2020). Psychological impact of COVID-19 pandemic in the Philippines. *Journal of Affective Disorders*, 277, 379–391. <https://doi.org/10.1016/j.jad.2020.08.043>
- Tikhomirov, O. (1988). *The Psychology of Thinking*. Moscow: Progress Publishers.
- Tsakiris, M. (2010). My body in the brain: A neurocognitive model of body-ownership. *Neuropsychologia*, 48, 703–712. <https://doi.org/10.1016/j.neuropsychologia.2009.09.034>
- Türkyılmaz, Ç. (2015). Hegel and Marx on the Self-Realization of Man. *Hegel-Jahrbuch*, 2015(1), 251–255. <https://doi.org/10.1515/hgjb-2015-0140>
- Van der Veer, R., & Valsiner, J. (1991). *Understanding Vygotsky: A quest for synthesis*. Blackwell Publishing.
- Vygotsky, L. S. (1978). *Mind in society: The development of higher psychological processes*. Harvard University Press.
- Vygotsky, L. S. (1986). *Thought and language*. MIT Press.
- Vygotsky, L. S. (1987). The psychology of Schizophrenia. *Soviet Psychology*, 26(1), 72–77. <https://doi.org/10.2753/RPO1061-0405260172>
- Vygotsky, L. (1997a). The historical meaning of the crisis in psychology: A methodological investigation. In R. W. Rieber & J. Wollock (Eds.), *The collected works of L.S. Vygotsky. Cognition and Language (a series in Psycholinguistics)*. (Vol. 3). Springer. https://doi.org/10.1007/978-1-4615-5893-4_17
- Vygotsky, L. (1997b). Mind, consciousness, the unconscious. In R. W. Rieber, & J. Wollock (Eds.), *The Collected Works of L.S. Vygotsky. Cognition and Language (A series in Psycholinguistics)* (R. Van Der Veer, Trans., Vol. 3, pp. 109–121). Springer.
- Wathelet, M., Duham, S., Vaiva, G., Baubet, T., Habran, E., Veerapa, E., & D'Hondt, F. (2020). Factors associated with mental health disorders among university students in France confined during the COVID-19 pandemic. *JAMA Network Open*, 3(10), e2025591–e2025591. <https://doi.org/10.1001/jamanetworkopen.2020.25591>
- Wittgenstein, L. (1958). *Preliminary studies for the "Philosophical investigations" generally known as the blue and brown books*. Blackwell.
- Woźniak, M. (2018). I" and "Me": The self in the context of consciousness. *Frontiers in Psychology*, 9(1656), 1–14. <https://doi.org/10.3389/fpsyg.2018.01656>
- Zaytseva, Y., Chan, R. C., Pöppel, E., & Heinz, A. (2015). Luria revisited: Cognitive research in schizophrenia, past implications and future challenges. *Philosophy Ethics and Humanities in Medicine*, 10(1), 1–11. <https://doi.org/10.1186/s13010-015-0026-9>

Springer Nature or its licensor (e.g. a society or other partner) holds exclusive rights to this article under a publishing agreement with the author(s) or other rightsholder(s); author self-archiving of the accepted manuscript version of this article is solely governed by the terms of such publishing agreement and applicable law.

Mohamad El Maouch Dr Mohamad El Maouch finished his Ph.D. in Psychology, from Northeast Normal University (China), and his internship at the Institute of Automation, Chinese Academy of Sciences (Beijing), he also did his Post-Doc at the Lebanese-American University (LAU) (Lebanon- Beirut). Currently, he works as a research fellow at Zhengzhou Normal University. He is focusing mainly on Psychology and AI. Additionally, he is engaged in research investigating sociohistorical aspects of neuropsychology, mental development, mental health, creativity, and personality, by following the guidelines of the cultural-historical activity theory (CHAT). He published conferences and journal papers including theoretical and empirical papers about various topics.

Zheng Jin Dr Zheng Jin graduated from Gyeongsang National University (the Republic of Korea-ROK), with Ph.D. in Human Factors. Currently, he is a Professor and the Director of the International Joint Laboratory of Psychological Science at Zhengzhou Normal University. He is also a collaborative researcher at the University of California at Davis since 2012. He has been responsible for the coordination of several funded scientific research projects at the National level. In addition to his research into Implicit Cognition, he has published many articles, book chapters, and conference papers on psycholinguistics and perception to understand Ecological Models that link perception with actions.