

This is a repository copy of *An AI tool for scaffolding complex thinking: challenges and solutions in developing an LLM prompt protocol suite*.

White Rose Research Online URL for this paper:

<https://eprints.whiterose.ac.uk/id/eprint/229362/>

Version: Published Version

Article:

Teixeira de Melo, Ana, Caves, Leo, Renault, Letícia et al. (4 more authors) (2025) An AI tool for scaffolding complex thinking: challenges and solutions in developing an LLM prompt protocol suite. *Cognition, Tehcnology & Work*. ISSN: 1435-5558

<https://doi.org/10.1007/s10111-025-00817-6>

Reuse

This article is distributed under the terms of the Creative Commons Attribution (CC BY) licence. This licence allows you to distribute, remix, tweak, and build upon the work, even commercially, as long as you credit the authors for the original work. More information and the full terms of the licence here:

<https://creativecommons.org/licenses/>

Takedown

If you consider content in White Rose Research Online to be in breach of UK law, please notify us by emailing eprints@whiterose.ac.uk including the URL of the record and the reason for the withdrawal request.



An AI tool for scaffolding complex thinking: challenges and solutions in developing an LLM prompt protocol suite

Ana Teixeira de Melo¹ · Letícia Renault² · Leo S. D. Caves³ · Philip Garnett⁴ · Paula Duarte Lopes⁵ · Raquel Ribeiro¹ · Filipe Santos¹

Received: 30 April 2025 / Accepted: 3 July 2025
© The Author(s) 2025

Abstract

This paper reports an exploratory study examining the interaction between a theoretical framework for Complex Thinking and AI (LLMs), in terms of its potentialities and constraints. The aim was to develop and conduct a preliminary pilot evaluation of a tool comprising a prompt protocol suite for use with an LLM, to scaffold Complex Thinking. The tool is designed for use by an individual or group in relation to a given Target System of Interest (i.e., a real-world system, a problem, or a concern), supporting the development of more complex understandings of such systems that can guide more effective and positive actions and decisions. We describe the process of developing a suite of prompt protocols for scaffolding particular properties of Complex Thinking and report on the outcomes of a pilot test evaluation with a set of users across different domains.

Keywords Complex Thinking · Large Language Models · Prompt protocols · Scaffolding

1 Introduction

1.1 Complex challenges requiring complex responses: Complex Thinking

Current global challenges have highlighted the inadequacy of the dominant modes of human thinking and the decision-making processes that address complex problems (Rogers et al. 2013; Wells 2013; Mancilla García et al. 2020). Worldwide, multiple risk processes interact at various scales and levels, in complex ways, driving social and ecological crises, often with unexpected but profound impacts on the lives of individuals and communities both locally and globally

(Renn 2017). The pressure to change our dominant stances and modes of thinking and to embrace the complexity of the challenges at hand is greater than ever (Morin 2005/1990). Building on the work of Ashby (1958), Caves & Melo (2018) proposed that the modes of thinking underlying decision-making and action in complex systems (e.g., interventions in social-ecological systems) should be congruent with the complexity of the systems they seek to understand and affect. Drawing on knowledge of complex system organisation, particularly in living and adaptive systems, Morin (2005/1990), Morin (2014) formulated guiding principles for a complex mode of thinking. There has been a growing recognition of the need to train and educate for complexity and to develop ways of operationalising some of those principles (Morin 2002; Ramírez-Montoya et al. 2021; Fielder-Ferrara 2010). Building on research on complex biological and social systems and Morin's proposal, Melo (2020) presented a theoretical framework to support the practice of Complex Thinking (CT). The notion of CT is not simply defined, as there are a variety of definitions and perspectives circulating in the literature which refer as much to the contents as to the processes of thinking (Melo 2020). More recently, there has been a growing interest in CT defined as a kind of higher-order thinking or as a core meta-competence, of critical importance for the 21st Century (Ramírez-Montoya, et al. 2021). While some of these authors refer to Morin's

✉ Ana Teixeira de Melo
anamelopsi@gmail.com

¹ Centre for Social Studies, University of Coimbra, Colégio de S. Jerónimo, Apartado 3087, 3000-995 Coimbra, Portugal

² Independent Scholar, Centre for Social Studies, University of Coimbra, Coimbra, Portugal

³ Independent Scholar, Vila Nova de Gaia, Portugal

⁴ School for Business and Society and the Institute for Safe Autonomy, University of York, York, UK

⁵ Centre for Social Studies, School of Economics, University of Coimbra, Coimbra, Portugal

thinking, most equate CT with the integration or combination of meta-cognitive capacities, critical thinking, creative thinking and systemic thinking (Baena-Rojas et al. 2022; Ramírez-Montoya, et al. 2021, 2022; Silva Pacheco & Iturra Herrera 2021), as well as reflexivity (Morales 2020). In this alignment, more studies are emerging regarding interventions to promote these skills (Ramírez-Montoya, et al. 2021; Olivo-Montañó et al. 2024), particularly those that explore the use of technologies (Patiño et al. 2023), especially in the context of higher-education, where attempts have also been made to develop evaluation scales (Tobón & Luna-Nemecio 2021).

In the framework proposed by Melo (2020), CT is grounded in a relational ontology and a view of cognition as embodied and enactive (Varela et al. 2016). It conceptualises both a process and an outcome of the coupling between a given Observer and their environment (mutually defined). As a process, CT attends to and recognises properties of complex systems but also enacts them in its own organisation, thereby *performing* complexity (e.g. leading to novel, surprising emergent ideas or possibilities for action). The more complex the thinking, as an outcome, the more it (i) affords a multiplicity of descriptions, explanations and anticipations, (ii) leads to emergent knowledge information that increases the coherence of the coupling between the observer and their targets of interest; (iii) expands the possibilities for effective actions for promoting, supporting or managing changes, leading to more positive results from the perspective of the widest variety of critical observers in a system and (iv) supports constructive interactions amongst them and their positive co-evolution (Melo 2020). The proposed framework lists 9 dimensions along with a set of 24 corresponding properties of complexity characterising the process and outcomes of CT: structural (structural variety and dimensionality, relationality); dynamic/process (recursiveness, dynamic processes, relativity/ambiguity/uncertainty); causal and explanatory (modes and finalities, historicity, complex circularity, emergence); dialogic (dualities and complementary pairs, trinities); observers' (multipositionality, reflexivity, intentionalities); developmental and adaptive (developmental-adaptive value, developmental evolvability); pragmatic (pragmatic value, pragmatic sustainability); ethical (ethical value) and aesthetical (aesthetical value) and narrative (differentiation and integration, identities, flexibility/openness). Although these properties, conceived as *thinking movements*, can be executed in isolation, with different levels of complexity, CT requires the coordination and interaction between them to create conditions for the emergence of abductive and emergent leaps which generate novel and meaningful information, for a particular context. This information may guide the actions and decision-making of the observer in relation to their Target Systems of Interest (TSOI) (SEBoK 2024; The Open University 2023) in

ecosystemically fit and effective ways. CT can also be seen as the contributions of an observer engaged in a complex coupling (or dance) with their environment(s), characterised by the unfolding of different dynamic configurations of properties through time (Melo and Renault 2025). These configurations vary in their capacity to generate creative and abductive leaps and produce new, relevant information. Many theoretical approaches to CT are restricted to what could be called “complexity thinking”, focusing more on the contents of the thinking than on the processes (Melo et al 2019). However, even proposals such as Morin's (2005/1990), which are focused on the process, do not elaborate on the importance nor the nature of the relation between thinking movements. The focus on the importance of the interaction between properties or movements of CT, namely in leading to abductive leaps, is a distinctive feature of the CT model proposed by Melo (2020). However, this framework requires further investigation into how and under what conditions these properties interact, and what strategies should support that interaction. Operationalising this framework with LLM support, even when focused on a subset of properties, prompts reflection on the nature of these thinking process interactions and may offer insights for future model development.

Theoretical frameworks on CT are critical for developing tools and strategies to support decision-making on complex problems. Such tools should guide CT implementation to expand individual and collective intelligence and cognitive capacities.

1.2 AI and scaffolding CT

Artificial Intelligence (AI) is often presented as a powerful route to address complex challenges and find adequate responses for the wider 'social good' (Manyika 2022). Cows et al. (2021, p. 114) state “designed well, AI technologies can foster the delivery of socially good outcomes with unprecedented scale and efficiency”. Recent advances in AI have opened opportunities and lowered barriers to managing large amounts of information in ways that may contribute to building complex knowledge. However, there are other possibilities for dealing with complex problems, such as the use of AI systems to promote (relatively more) complex modes of thinking in humans, or the assembly of hybrid systems of collective intelligences capable of performing it (Cui & Yasseri 2024; Hemmer et al. 2021; Jarrahi et al. 2022; Kamar 2016). The notion of Zone of Proximal Development (ZPD), as proposed by Vygotsky (1978), entailing a relational ontology (Stetsenko 2017), refers to the idea that we can exercise cognitive capacities that are beyond our usual performance when in the context of a relation with a more competent other, that *scaffolds* our performance. The re-visitation of this ZPD, through the scaffolding, may lead

to an expansion of that entity's capacities to the point where they are transformed and a new ZPD is created. By maintaining coupling with an entity, and gauging their ZPD, a scaffolder may affect a process of continuous development.

Recent advances in the domain of AI have offered opportunities for assembling systems of collective intelligence (Cui & Yasseri 2024), which expand the possibilities for thinking and possibility thinking (Beghetto 2023). While a deeper understanding of cognitive processes can contribute to the development of more sophisticated forms of AI (Marcus & Davis 2021) current tools may already have the potential to operate as minimally effective "partners in cognition" (Salomon et al. 1991) and to adopt a scaffolding role for human cognition (Jong 2005).

The role of a scaffolder is just one among many roles that an AI can assume in the coupling with humans (e.g. coach, assistant, companion) (Cui & Yasseri 2024). As a cognitive scaffolding partner, we would expect an AI to be able to guide the thinking of a given Observer in relation to their Target System of Interest (TSol) and to assist them in operating in their ZPD (Sætra 2022), but not to think "for them". Under certain circumstances, when AI-human coupling is sufficiently complex, co-augmentation of intelligence may emerge, producing capabilities and outcomes that exceed the individual contributions of either agent (Melo 2022).

Building on the work of Wood et al. (1976), Sætra (2022) highlights six characteristics of effective scaffolding: (i) motivation, (ii) direction maintenance, (iii) frustration control, (iv) simplification, (v) marking critical features, and (vi) demonstration. Holton and Clarke (2006) had previously expanded the notion of scaffolding, from Wood et al. (1976, cit in Holton & Clarke 2006) and Vygotsky's ZPD (1978), with a conceptualisation framework that includes the following dimensions: (i) scaffolding agency (expert, reciprocal, or self-scaffolding); (ii) scaffolding domain (conceptual and heuristic); (iii) equating self-scaffolding with metacognition; (iv) zones of scaffolding which combine the previous criteria to compose a gradation of types of scaffolding situations (e.g. moving from conceptual to heuristic, from expert to self-scaffolding). Holton and Clarke (2006) review other proposals of sets of scaffolding questions for supporting metacognition (which they view as a form of self-scaffolding) that can be used across these zones. The idea that cognition can be scaffolded through questions that would help the cognitive agent move beyond their current understanding points to ways in which AI agents can be mobilised into scaffolding roles to promote CT. Forms of dialogical scaffolding that utilise argumentation-based reasoning have also been proposed as a way to enhance the scope and quality of joint reasoning of humans and AI and to increase value alignment (Mogdil 2018). Question-based scaffolding is particularly well-suited for LLM-based AI that supports natural language interactions.

In exploring the potential role of AI in scaffolding human learning, Sætra appeals to work on Intelligent Tutorial Systems (ITS) by Brusilovsky (1999, cit in., Sætra 2022) which are organised in 4 modules: (i) a *teaching module* (with domain facts, rules and knowledge); (ii) a *learner model* (with learner skills, progress and general profile) and a (iii) *tutoring strategy* (theory of learning and tutoring), and an (iv) *interface*. The author proposes that a specialized adaptive AI expert scaffolding ITS can be similarly organized, with or without human expert assistance.

1.3 Large language models and prompt engineering

Generative LLMs have opened new possibilities for Human-AI interaction towards the assemblage of systems of augmented intelligence. LLMs, arising from the field of machine learning, are generative AI systems that are built upon underlying foundation models that are pre-trained to learn the relationships between tokens (typically fragments/roots of words) extracted from multiple extensive text corpora. The distinctive aspect of LLMs is the vast number of parameters (such as weights, biases, embedding vectors) that specify them: e.g. OpenAI's GPT 3.5 model is estimated to have around 20 billion parameters. LLMs can also have additional layers of training that fine-tune their output (text) in response to input (text aka prompts). This training can include evaluation of: completion of specific tasks, ability to follow instructions, and importantly, the degree to which responses stay within the bounds of a defined set of norms and values (aka alignment). The performance of LLMs is judged through the quality and relevance of their responses across various scenarios (for more information on LLMs see Naveed et al 2023). ChatGPT's launch in November 2022 crossed critical thresholds of performance and accessibility, generating massive publicity and triggering an explosion in developers, applications, and users that accelerated the AI boom. This success led to numerous popular LLMs becoming available through free or subscription models and increased AI integration across technology platforms.

LLMs generate outputs in response to prompts. White et al. (2023) define a prompt as a set of instructions that programs an LLM by customizing and enhancing its capabilities, influencing subsequent interactions and outputs by providing specific rules and guidelines for the conversation. Specifically, a prompt "sets the context for the conversation and tells the LLM what information is important and what the desired output form and content should be"; prompt engineering is therefore the way in which "LLMs are programmed" (White et al. 2023, p.1). There are four general elements of a prompt: (i) an *instruction*, (ii) a *context* (to steer better responses), (iii) *input data or question* and (iv) an *output indicator* (type or format) (Saravia 2022). There

are also a variety of approaches to prompting, aimed at optimising the interactions with the AI, usually emphasising iterative processes (Velásquez-Henao et al. 2023; Ray 2023; Sarrion 2023), and highlighting aspects such as detailed descriptions, context specification and step-by-step reasoning approaches (Ye et al. 2023; Ray 2023).

Some patterns have been identified that may be more effective for particular LLMs and can be adapted to different contexts (White et al. 2023). However, aiming for specificity and accuracy can lead to neutrality and less meaningful outputs, prompting some authors to adopt an alternative hermeneutic approach focused on meaning (Henrickson & Meroño-Peñuela 2023). Although sharing some similarities with programming, prompt engineering is more of an art or craft than traditional programming (OpenAI 2023; Google 2024; Zamfirescu-Pereira et al. 2023). There are general rules of thumb to guide the formulation of prompts that can be applied across most LLMs, but each LLM has its own features, or ‘personalities’, requiring specific adjustments of the prompting process.

The creativity level of LLM responses to prompts can be altered by changing the model's ‘temperature’: lower temperatures produce more predictable output, while higher temperatures generate more creative and varied but potentially less coherent responses (OpenAI 2023). A balance must be found between allowing the LLM creative freedom and keeping responses on-topic. Higher creativity levels increase the likelihood of ‘hallucinations’ (OpenAI 2023) – LLM-generated nonsense. The significance of this problem depends on the specific task. In our context, we wanted to constrain the LLM within the CT framework while allowing creativity in specific operations (e.g., metaphor generation). We developed protocols that achieved this balance at the default LLM temperature (see Sect. 2.3.1).

1.4 Aims and objectives

This paper describes an exploratory study that aimed to examine the interaction between a specific CT framework (Melo 2020) and generative AI based on LLMs, to develop a tool for scaffolding CT. The tool was designed to be used by individuals or groups for addressing real-world systems, problems, or concerns, to develop more complex understandings of such systems to guide more effective actions and decisions.

The study had three objectives:

- (i) to develop and conduct a preliminary evaluation of prompt protocols for scaffolding CT in individuals or groups;
- (ii) to identify the possibilities and limitations of integrating a CT framework with LLMs;
- (iii) to conduct a preliminary evaluation of the user experience with the tool developed.

We present the process and pilot evaluation results for a tool, comprising a suite of prompt protocols for an LLM, designed to scaffold specific properties of CT.

2 Method

Given the exploratory nature of our research questions and goals we adopted a general qualitative exploratory multiple case design (Yin 2018; Mills et al 2010; Saldanã, 2016), involving an evolving and adaptive developmental evaluation approach to guide the development of innovation (Patton 2011) and an iterative design process (Nielsen 1993).

2.1 Participants

During development, two participant groups—both internal and external to the team—were involved in pilot testing and evaluating the tool (LLM + prompt protocols suite). Table 1 summarises the participant characteristics and their engagement.

Two team members (*int.dev#1* and *int.dev#2*) were deeply involved in developing the prompt protocols and conducting preliminary tests during the initial phases. One authored the theoretical framework, and both were experienced with the framework and CT evaluation methods. Another team member (*int.dev#3*) collaborated on earlier protocol development but did not participate as a test user. Two additional team members (*int.tech#1* and *int.tech#2*) assisted with internal tests and addressed technical issues but did not participate as protocol users. Later, two other team members (*int.user#1* and *int.user#2*) who had not been involved in protocol development acted as users in pilot case studies 1 and 2, testing more robust protocol versions.

Case studies 5 and 6 included a second subset of external participants with very applied focuses. These participants corresponded to two community-based teams of practitioners who had many years of working together as a team. These teams are experienced in conducting assessments and interventions of families with at-risk children. They both had prior training in (multi)systemic intervention approaches and in CT, and were familiar with the process of case conceptualisation. They participated in the tests as a team, engaging with the tool collectively. All participants in the study provided written informed consent regarding the conditions for their participation and data collection.

Table 1 Characterisation of the participants of the study and their engagement

Participant code	Case study	Type ^a	Gender ^b	Level of Education / Training ^c	Domain /Area of intervention	AI experience ^d , Research Domain used as TSoI ^e ?
<i>int.dev#1</i>	-	INT	F	PhD	Psychology / Community and Family interventions	Minimal
<i>int.dev#2</i>	-	INT	F	PhD	Psychology	Minimal
<i>int.dev#3</i>	-	INT	F	PhD	International Relations	None
<i>int.tech#1</i>	-	INT	M	PhD	Computational biophysics / Complex systems	Experienced
<i>int.tech#2</i>	-	INT	M	PhD	Complex systems/Organisations	Experienced
<i>int.user#1</i>	1	INT	M	PhD	Sociologist researching the execution of security measures for mentally disordered offenders	Experienced; Yes
<i>int.user#2</i>	2	INT	F	PhD	Social psychologist researching housing issues and inequalities	None; Yes
<i>ext.user#1</i>	3	EXT	M	M	Futures research	Experienced
<i>ext.user#2</i>	4	EXT	F	M	Applied research on social-ecological systems	None
<i>ext.user#3</i>	5	EXT	F (×2)	UG; PG; M	Social work; Psychology	None
<i>ext.user#4</i>	6	EXT	F (×2)	UG; PG; M	Social education; Social Worker	None

^aINT, internal team member; EXT, external participant^bM, Male; F, Female^cUG, undergraduate degree; PG, postgraduate training; M, Masters^dLevel of prior experience with AI tools^eWas the participant's research domain used as the TSoI in testing? (only affirmative indicated)

2.1.1 Ethical Approval

This study obtained ethical approval (REF 90/24) from the Ethics Committee of the [blinded for review].

2.2 LLMs as a platform for the scaffolding tool

For our purposes, we explored using LLMs as a platform for developing the scaffolding tool. Given project time constraints and available resources, we used a text prompting approach through web interfaces rather than exploring APIs. We wanted to constrain LLM behaviour without the use of specific technical parameters (such as temperature settings or pre-defined personas with different creativity levels) by providing protocols that 'encode' aspects of the CT framework. These protocols shape LLM responses to human prompts, enabling collaborative human-AI thinking processes.

During protocol development, we experimented with different LLM platforms (at pro/paid subscription levels): ChatGPT 4/4o (OpenAI), Claude Sonnet 3.5 (Anthropic), and Sonar (Perplexity). We also tested other platforms (e.g., Google Gemini, a local Ollama instance) to a lesser extent. We began with ChatGPT 4o, then conducted parallel comparisons between ChatGPT and Perplexity using its default model, and later compared Claude Sonnet 3.5 with ChatGPT 4o. The final protocols were fine-tuned for Claude, as described below.

2.3 Development and evaluation of the protocols

The development of prototypes of prompt protocols for scaffolding CT with LLMs proceeded through an iterative process comprising a set of key phases, described below. The process was iterative both within and between phases. We implemented and conducted preliminary checks and tests on the performance of the protocols during the development phases and several rounds of refinements after the pilot tests with internal and external users. We conducted a type of developmental evaluation (Patton 2011), an evolving process that unfolds coupled to the development of the protocols, thereby informing the process and guiding the necessary adjustments. The protocol evaluation, addressing our first and third objectives, focused on two areas: 1) the degree to which the tool adhered to the CT framework and fulfilled scaffolding functions, and 2) factors related to user experience.

Regarding the first focus, we wanted to ensure that, during the interaction with the user, the AI was capable of responding in accordance with the theoretical framework and with a set of requirements defined as critical (see Table 2). A fundamental step in developing the protocols involved creating and calibrating an evaluation framework for CT, which was embedded in the prompt protocols for the AI to evaluate users' relative thinking complexity. This framework guided the AI in coupling with users and adjusting its responses based on that evaluation (in essence, the tool gauges the user's ZPD). We also wanted to ensure that, with different

Table 2 CT Framework Requirements: Challenges for Human Users and AI Tools, and Solutions Implemented

Requirements	Challenges to humans	Challenges related to AI (LLM) tools	How challenge was addressed
CT requires the performance of various types of thinking movements of different degrees of difficulty	Difficulty of some movements (e.g. restrictions of memory; multidimensionality of information); Limited repertoires of practices	Difficulty of some movements to an AI (e.g. intentionality, pragmatic, aesthetic, ethical values) (cf. alignment problem); Lack of social and cultural sensitivity.	AI is prompted to support thinking movements that are possible to scaffold and prompted to guide the user to make external ethical, aesthetic, social and cultural checks
CT requires the recursive interaction between different thinking movements and their outcomes	Difficulty of keeping awareness of various thinking movements to choreograph configuration	Difficulty attuning to intentionality and purpose; restricted modes of coupling with the observer (mostly verbal I); “Lack of intuition” (e.g. cue in); Need for a “meta-tool” to orchestrate the coordination of scaffolding modules for different properties	The AI guides the user in performing a target thinking movement/property. Other properties are embedded in the scaffolding of the target property. No meta-tool is currently available.
Information emergent from relations and relational movements requires exploration	Limited memory and mapping capacities for visualising and navigating the relational structure of information	Limited memory and mapping capacities for visualising and navigating the relational structure of information	AI invites users to capture relational information through metaphors and offers visualisations. AI creates summaries and basic visual maps. Relational mapping still limited
The (configurations of) thinking movements (processes) need to be dynamically traced along contents	Limitations of meta-cognitive capacities to track process along contents	Needs to map various thinking movements and their coordination and the nuances of their complexity along the unfolding of contents. Needs specific “training” and a reference database	Not currently addressed
Ongoing reflexivity (in self and hetero scaffolding);	Difficulties with awareness of “blindspots”, personal constraints, habits and how they shape the ongoing thinking;	Absence of “true” reflexivity (lacks self-awareness)	AI can prompt and induce reflexivity and invite the user to conduct activities that induce reflexivity, namely involving other reflexive agents
It is necessary to allow for a (recursive, embodied, active) manipulation, mapping, updating and retrieval of complex information (e.g. relational, multi-level, multi-temporal scales)	Limited memory and capacity to hold complex information	Memory and limitations with interface (in terms of allowing physical manipulation of information)	AI prompts the user to perform active embodied activities, outside of the tool, to physically and metaphorically manipulate and explore the information Claude’s Artifacts were used for various outputs, namely visual, allowing the user to track, interact with and manipulate the dialogue
Emergence (new hypothesis, ideas) requires experiments with multiple alternative thinking trajectories s. Path-dependence and independence both need to be explored	Non-linearity and randomness of thinking movements are difficult to manage (path-independent)	AI needs to forget a pathway to open new ones, but also to remember to guide the user in exploring its contingencies and to find patterns	Not currently addressed
‘Fitness’ (e.g. pragmatic, ethical value and integrity and accuracy) of the thinking and information and ‘good enough’ complexity	Limited perspectives; preconceptions, prejudices, stereotypes. Double reflexive effort for humans: to keep themselves and the AI in check	The problem of association; “hallucination”, speculative, prejudices, stereotypes; Absence of pragmatic, aesthetic, ethical values; Inability of directly coupling the target system in terms of values	AI is prompted to have the human perform checks with external sources

Table 2 (continued)

Requirements	Challenges to humans	Challenges related to AI (LLM) tools	How challenge was addressed
The scaffolding needs to generate coherence and (sufficient) difference in the coupling between entities or with the TSol	The complexity of the human will affect their coupling with the framework and their metacognitive capacities for an ongoing nuanced evaluation to guide the coupling with another Human or TSol	Difficulties in the coupling of the AI for an ongoing performance evaluation that guides the coupling with another Human or TSol are a challenge Difficulties of “attuning” (multiple types of coupling) to the observer and their domain and the social-cultural specificity of the context variety of ways CT movements can be performed; Asymmetrical coupling AI-human	Not directly addressed. Recommendations for using the protocols with a Human scaffolder familiar with and with expertise in the framework and with sufficient complexity to couple both with the human user and the AI

kinds of problems and inputs, the AI would follow the protocols sufficiently, providing responses informed by the chosen CT theoretical framework and avoiding the issue of LLM “hallucinations”. This evaluation was based primarily on AI chat transcript analysis and user discussions. Different protocol versions used in user-AI interactions and their corresponding chat transcripts were systematically compared to evaluate the effects of prompt variations and protocol structure changes.

Regarding the second focus on user experience, we were guided by questions such as: (i) What characterises the user experience in the interaction with the AI? To what extent is the interaction perceived as “easy”?; (ii) To what extent is it clear for the user how to proceed during the interaction?; (iii) To what extent do the users experience the interaction as valuable or as bringing forth something new or relevant to their understanding and coupling with their Target Systems of Interest?; (iv) To what extent do the users feel that the AI is guiding and stimulating their thinking (as opposed to thinking “for them”) and well coupled with it?.

2.3.1 Phase 1: Requirements and aims of the scaffolding process

In this phase, we analysed the requirements set up by the theoretical framework on CT and clarified the aim of the scaffolding process.

Requirements and challenges of the implementation of a CT Framework. The theoretical framework guiding this study (Melo 2020) encompasses a particular set of assumptions about the nature of cognition as an embodied, enactive, embedded and extended process (Newen et al. 2018) and its underlying (predominantly constructivist) ontology and epistemology. Cognition is understood as an activity emergent from the coupling of a particular observer with its environment. The framework also proposes a very particular definition of CT. The design of any tool or strategy to scaffold CT needed to be congruent with the theoretical assumptions and foundations of the framework. The theoretical framework then sets up specific requirements which pose particular challenges, in particular for the development of an AI-based tool. Table 2 lists some of the core requirements derived from the framework that guided the decisions and problem-solving efforts of the team, along with the challenges identified for both humans and the AI, and how we addressed the challenges. We did not consider finding solutions for human and AI separately. Instead, we considered these challenges together and their interactions as developmental constraints and designed approximate, temporary responses to bootstrap the process. Table 2 includes notes on requirements for which we could not find suitable or partial solutions. It is important to acknowledge that these requirements stem from the framework's philosophical foundations

and that different conceptions of cognition and CT conceptualisations would likely have led to different types of AI (LLM) interaction constraints.

In the future, the pragmatic effects of these requirements and of the solutions proposed should be thoroughly investigated under the scope of a mixed-methods approach (Varga 2018) capable of capturing their multidimensional and multilevel implications of the strategies adopted along the nuances of the users' rich experiences and the variations associated with different strategies.

Aims of the scaffolding process. Our goal was to develop a set of protocols for AI (LLM) tools to support individuals (or groups) in performing CT in relation to a given TSoI (a real-world system, a problem, a concern) to develop more complex understandings of such systems that could guide more effective and positive actions and decisions. Our focus was not to have an AI tool providing answers or solutions to a problem, but to engage with and guide a user, within the principles of the CT framework, in their own process of thinking to increase its complexity. Hence, the target thinking complexity is located in the coupling relation between the individual and their TSoI, although our aim was also for the the AI interaction process to be sufficiently complex to generate emergent insights and ideas. Our aim for the scaffolding tool was to stay relatively coupled to the user's thinking, working from their existing complexity level and then providing support for more complex thinking movements. The tool should provide scaffolding that introduces a degree of novelty but not excessively. Any suggestions should support users in moving towards relatively more complex thinking, but not "think for" the user. This created a strong requirement for a robust CT evaluation framework that would set up key constraints for LLM tool operation and allow it to make proposals based on the complexity of the user response.

Critical guiding criteria in developing the scaffolding functions of the prompt protocols. The development of the protocols was guided by the chosen CT framework and, therefore, the aim was for the tool to operate within that framework. We also aimed to develop protocols that allowed the AI to perform scaffolding functions. Hence, the notion of scaffolding provided another set of requirements. Based on the literature, we considered that the AI should, at least be able to (i) couple with the user at the level of the current organisation of their thinking and, from there, (ii) introduce some relative (but not excessive) variation to support the user in moving through their ZPD. We were sensitised to these demands as we were systematically evaluating the different versions of the protocols through the results of using the tool. We noticed that some criteria were achieved but not others. As we refined the protocols, we aimed to balance these different dimensions and became more aware, through inductive analysis of the outputs, of the criteria we were

trying to attune and balance. We identified three main criteria which we called: *Boundaries*, *Attunement* and *Variation* (see Fig. 1). We intentionally looked at these criteria and worked to balance them in relation to each other: Boundaries relate to the requirement to stay within prescribed (or desired) *bounds*, e.g. confining the AI responses to be within the CT framework itself by taking its evaluation framework as a reference. Essentially, this meant seeking behaviour that was *keeping within the scope and context* of the phases, steps and modes of the protocols. Another distinctive requirement was to maintain the *focus of attention* on the *processes* of the user(s) thinking rather than on the content. An additional ethical consideration was to instruct the tool to refrain from making any decisions (for the user) and offering any suggestions for interventions or decisions in relation to the TSoI.

Boundaries were of greatest importance and considered foundational—they had the greatest weight in our evaluations. Attunement relates to *adapting* to users' contexts. Examples include adapting both the selection and modification of generic scaffolding questions to be appropriate to users' particular contexts and adapting protocols (e.g., choice of moves within/between modes) according to user responses. Another aspect is *engaging* with users, including adopting a dialogical style that adapts to users' own ways of expressing themselves and stimulating curiosity and motivation for engagement. Attunement was supported by Boundaries through developing a CT evaluation framework embedded within protocols that guided the AI in evaluating users' thinking complexity and supporting response adjustment. Variation relates to the requirement that the tool adds or creates information as "differences that make a difference" (Bateson 1979, p. 99). This can take many forms, such as adding *organisation* or *nuance* to existing information. An important CT framework aspect was for the tool to *suggest embodied activities* that go beyond the more propositional types of knowledge elicited by the principally text-based interaction mode. These additions are characterised by their *novelty* to users and the experience of *surprise*. A combination of Attunement and Variation was critical for the tool to *explore and flesh out metaphors* offered by users, offering nuanced perspectives whilst staying within the scope of user-constructed meaning. With the other two criteria being satisfactorily met, Variation was also very important in scaffolding. The AI should not be too rigid nor too loose (and unbounded) in its responses and should expand users' repertoires of thinking movements.

Variations in prompt design, their configuration within protocols, and their relations to other protocols in the suite resulted in changes in the expression of the guiding criteria for evaluating scaffolding performance. The approach to balancing criteria was one of satisficing rather than optimising.

Choice of LLM. Our choice of LLM was informed by comparing their performance on the three criteria described

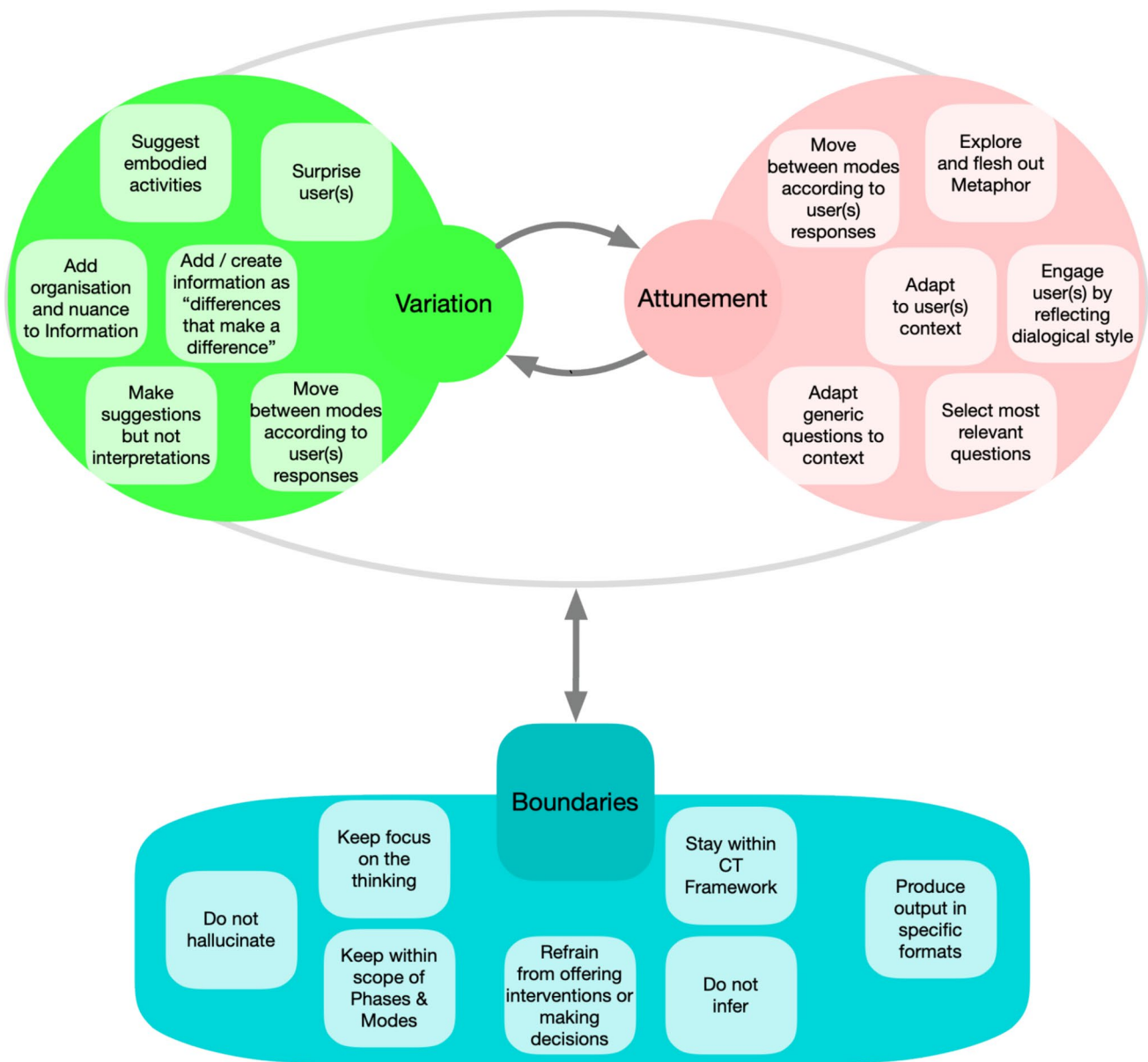


Fig. 1 Critical criteria guiding prompt development to achieve protocol balance

in the previous section. We found it much more difficult to achieve balance between these criteria with some models than others. Although not intended as a comprehensive assessment of their capabilities, we offer the following reflections on our experiences and the rationale for our choice of platform.

ChatGPT appeared eager and impetuous, tending to problematise prompts and readily offering "solutions," sometimes in a rather "lazy" manner, neglecting detail and "averaging" information. ChatGPT would also lose track of the protocol fairly quickly (losing boundaries), "running after" some theme that emerged, leading to collapse of the thinking process structure. Sonar was more structured

but appeared rather rigid and perhaps boring. However, it would adhere to the protocol quite closely (staying within boundaries) but without offering much creativity or insight (low variation).

Claude had the best balance between keeping close enough to the protocols (within boundaries) that it could facilitate the thinking process whilst also having sufficient creativity in its responses to engage with human users naturally and remain coupled with their thinking (variation with attunement), guided by the evaluation framework. For example, it incorporated aspects of the prompts into its responses (e.g., if the input style was more narrative, it would tend to mirror that in responses), which led to more

engaging interactions. The team as a whole found Claude to be an engaging, respectful, and "insightful" dialogical partner.

With the introduction of Artifacts (persistent, editable, versionable documents such as text and images), Claude met some of our technical requirements naturally, becoming the primary platform for developing the scaffolding tool (LLM + protocol suite). A benefit of the artifacts is that they can be accessed, visualised, and edited in parallel to the main dialogue (chat).

After identifying Claude as the model with the best fit for our intentions and best performance, the protocols continued to be developed, refined, and fine-tuned for this model in particular. The results described below are from the tool using the Claude 3.5 Sonnet model only.

2.3.2 Phase 2: Operationalising the theoretical framework

The adopted framework for CT proposes 9 dimensions and 24 properties of the thinking. It also postulates that the (non-linear) interaction between the thinking movements associated with the different properties is critical for the complexity of the thinking as an outcome and for creative and abductive emergence. Given the short duration of this exploratory project and the complexity of the framework, we targeted only a small subset of properties. We chose to focus on the dimension of *Structural Complexity of the Thinking* and the corresponding properties of *Structural Variety and Dimensionality* and of *Relationality*, given their centrality, according to the framework, to generating (sufficiently complex) *contents* of the thinking upon which other thinking movements are performed. A full implementation of a scaffolding system for CT would require designing specific protocol modules for scaffolding each property as well as meta-modules to guide the orchestration and integration of movements associated with the non-linear and recursive interactions between different properties. Such an effort was not feasible within the scope of the project; our initial aim was to establish the feasibility of using LLMs to scaffold CT. Nevertheless, since more CT is predicated on interactions between a variety of different thinking movements, it was necessary to ensure some degree of interaction or modulation of one property by another. Therefore, for each module we implemented, alongside questions and activities focusing on the target CT property, we created lists of *embedding modulating questions* to be used in combination with questions or activities focused on scaffolding the target property. These questions, although still focused on the target property, do so in relation to other dimensions such as the *Complexity of the Observer* and the properties of *Multipositioning*, *Reflexivity*, and *Intentionalities*.

2.3.3 Phase 3. Development of an evaluation framework for CT

As mentioned above, there was a strong requirement for setting boundary constraints. A robust evaluation framework for CT should establish the criteria to evaluate the current level of complexity of the expressed thinking of the user(s). This phase of work built upon experience of operationalising the CT framework for the development of a visual mapping method for CT, *Complexigraphy* (Melo & Renault 2023, Melo and Renault 2025). The aim was to identify the core distinctive features of the performance of each property, the definition of different levels of complexity, and evaluation indicators. For example, we developed a detailed evaluation of the sub-property of *multidimensionality and variety of contents of the thinking*, which considers a high number of content descriptive dimensions. Based on the selected theoretical framework (Melo 2020) and previous works (Caves & Melo 2018) we determined the dimensions that were necessary for a minimally complex description of a TSoI: (i) the *internal organisation of the system*, (ii) its *environment*, (iii) the *observer* and (iv) the *coupling between these dimensions*. Each of these dimensions had their own sub-dimensions. The evaluation required attending to a variety of sub-properties and indicators simultaneously and in relation to each other.

The evaluation framework for CT, when implemented, would allow the tool to adapt its responses in relation to the current level of complexity of the users' thinking. Several versions of this evaluation framework were created, and refined, alongside the development and calibration of the evaluation protocol to be used by the LLM.

2.3.4 Phase 4. Development and calibration of the evaluation protocol for CT

Several iterations were required in the development cycle of the prompt protocol for the evaluation of the complexity of the thinking for the dimension of *Structural Complexity*. We started with the evaluation of the property *Structural Variety and Dimensionality* and then moved onto the property of *Relationality*.

Protocol development. We started to implement the evaluation framework for CT by trying to provide the LLM with a detailed description of the properties, levels and indicators of complexity within the CT framework that could be used to evaluate a given text or case (initially we used some selected short case studies/reports across domains, though see subsection on Calibration below). At that time, supplying data to the LLMs via spreadsheet resulted in many errors that were never satisfactorily resolved. An alternative approach using documents with hyperlinks pointing to the definition of each property and indicators also failed. Ultimately, we

adopted a detailed step-by-step narrative approach to deliver the property descriptions and the evaluation protocol. It was critical to have the tool systematically track and accurately consider all dimensions and sub-dimensions since the evaluation of all other properties of CT is built and evaluated “on top” of this basic content structure. When provided with the protocol, the tool appeared to try to synthesise and provide only an overall evaluation of properties, and to ignore lower-level sub-properties or indicators, or to focus too much on some details, while ignoring others. Worse, it might also rely on its own interpretation of the properties. It was these kinds of challenges that shaped the development of our prompting strategies, including the structuring of the protocol documents and disambiguation of references to critical text via unique identifiers (summarised in Sect. 2.4.1). In later protocol development phases, we started to test recursive references to the evaluation criteria and to critical pieces of information, to ensure they were “carried over” between stages of the evaluations within a property, and between properties, and kept in consideration throughout the dialogue. This kind of recursive loop allowed us to stabilise the protocols and to ensure that the responses of the tool were grounded in the CT framework and attuned to the user(s) context.

Calibration of the evaluation of CT protocols. To support calibration of the protocols for the evaluation of CT, for each of the target properties we used the LLM to develop a narrative about a fictitious scenario where an Observer is reporting the problems and concerns of a given community via a descriptive narrative that intends to convey a complex understanding of the situation. We initiated the development of this scenario on ChatGPT 4o. The imaginary community was called *Rivertown* by the tool. We adopted a stepwise process whereby we instructed the model to add details to an initial narrative related to each of the sub-properties and indicators. We started with the development of a *high* complexity narrative and then developed relatively *moderate* and *low* complexity narratives from it. We conducted several tests of the evaluation protocol using the three *Rivertown* scenarios. We compared different LLMs and experimented with different approaches for the prompts.

2.3.5 Phase 5. Organisation of prompt protocols for different scaffolding modes

To attend to CT framework requirements, the protocols needed to include different scaffolding modes. We started developing two core modes: *dialogical scaffolding* and *evaluation*. The former was chosen to emulate dialogical interaction between an individual and a human scaffolder via the LLM chat interface. The latter was critical for both setting boundaries for LLM operation within the CT theoretical framework and evaluating users' thinking complexity. This evaluation mode's role is to integrate the evaluation

framework for CT (see Sect. 2.3.3). This evaluation is used across all protocols to adjust interactions accordingly.

The protocols underwent testing to facilitate progressive refinements and fine-tuning. However, these modes alone were insufficient to satisfy CT framework requirements (see Table 2), particularly regarding the enactive and embodied dimensions of thinking and the need to organise and synthesise information for user manipulation, reflection, and action. As the dialogical scaffolding and evaluation modes were performing minimally well, we expanded prompting to cover other scaffolding modes and address additional CT framework requirements, integrating them into the protocols.

The role and nature of these new modes truly emerged from the LLM interaction process through our evaluation of response limitations and their effects. For each mode, we also explored what artifacts could be created (see Sect. 2.3.1 on model choice) that could enhance the mode's scaffolding effect and address framework requirements regarding visualisation and embodied, enactive information manipulation. We went through several iterations refining prompts for each artifact, identifying technical limitations, and attempting to design ways of overcoming them. Some limitations were not fully addressed.

2.3.6 Phase 6. Composing prototype scaffolding modules for the target properties

After having developed prompts for the different scaffolding modes, we faced the challenge of deciding how to best integrate them into a single protocol, composing a module for each target property of the CT framework that would guide the whole of the scaffolding process. We had to decide how to sequence them, how much relevance to give to each and how much recursion and iterations to build within and between the different modes. The next phases were critical and we used the pilot tests to refine this process and to experiment with different forms and sequences of the prompts to generate different trajectories through the different modes.

2.3.7 Phase 7. Prototyping pilot testing with internal users

Once we achieved what we considered to be a minimally robust protocol we conducted more systematic pilot tests of the protocols. Case studies 1 and 2 included two internal users (*int.user#1* and *int.user#2*) who tested the protocols with different intentionalities and target problems, under supervision. This allowed us to identify additional limitations of the current approach and to introduce additional nuances in the modes of the protocols. There were two independent sessions conducted with each user. Each session lasted approximately 90 min. For subsequent analysis, these sessions were audio and video recorded and the logs of the chat were extracted from the LLM tool.

2.3.8 Phase 8. Prototyping pilot testing with external user pilot tests

Additional pilot tests (case studies 3 to 6) were conducted with external users in the presence of the team (one session for case studies 3 and 4 and two sessions for case studies 5 and 6). During the sessions the team provided support, sometimes intentionally nudging the user to try particular steps or approaches to test the robustness of the protocols. The team also recorded the user's most salient reactions and reports of their experience. Each session lasted approximately between 120 and 180 min. For subsequent analysis, these sessions were audio and video recorded and the logs of the chat were extracted from the LLM tool. At the end of each session there was a short interview to explore the participant's experience. Additionally, the participating teams completed a post-session reflection commenting on their session log.

2.3.9 Phase 9. Final refinement of the protocols

Based on the previous phase, opportunities for further improvements of the protocols were also identified, leading to the fine tuning of the protocols and the introduction of minor adjustments.

2.4 Technical aspects of the tool development process

2.4.1 Features of the adopted prompting strategy

We highlight the following aspects of our approach to developing the prompts that make up the evaluation and scaffolding protocols:

1. Use of documents with hierarchical structure, containing: (i) an *introduction* and (ii) definition of the *role* of the AI, followed by (iii) general *guidance*, which sets up the contexts, and details the different *modes* of scaffolding, corresponding artifacts and what is expected as an outcome and by the AI, (iv) instructions hierarchically organised into *phases* and *steps*, each calling upon different modes of scaffolding;
2. Adopting of an algorithmic approach with procedures and branching (e.g. "If-Then" conditional branching);
3. Use of identifiers (or tags) to ensure disambiguation of references within and between modules to text describing the context, phases and dimensions of the thinking. Identifiers were denoted by text between square brackets e.g. [DSCAFF MODE] (NB. any convention

would likely suffice e.g. hash tags, or HTML-like angle brackets etc.);

4. Adopting a hierarchical naming scheme for identifiers that reflects the structure i.e. dimensions and properties of the CT framework e.g. [SC.SVD.MD.O.MOD.OC.MP.FMODQ1]. This was applied to all the questions allowing us to evaluate more easily how they were being used by the tools;
5. The interaction (between the tool and the user) is framed by a "General First Instruction" document, which introduces the CT framework and defines CT and the organisation of the framework into dimensions and properties;
6. Non-linearity: the tool guides thinking iterations following the different steps sequentially, but also is instructed to give the user choices on how to proceed and which modes of scaffolding to engage with; depending on the responses some steps may be skipped;
7. Recursion: some choices may result in a re-entry into previous steps to deepen the scaffolding, forming recursive loops. This aligns with the importance of reflexivity in the CT framework;
8. Provision of guidance ([GUIDANCE]) to establish context for the LLM to try to minimise alternative "interpretations" or operating outside of the CT framework; Each phase has a general description of what is to be achieved
9. Use of precise instructions within the prompts to constrain responses and to try to define boundary conditions for the LLM to "stay within the framework"
10. Explicit reference through identifiers to particular documents/images that are generated/refined through the process (e.g. Claude's Artifacts).

These elements work together to create a structured system of internal references (e.g. to guidance instructions or documents such as those detailing the evaluation framework and its dimensions, properties sub-properties and indicators, along a detailed description of modes of scaffolding and the role of AI) to ensure (or at least promote) a "tight(er)" and disambiguated protocol such that the AI's suggestions and guidance were properly framed by the theoretical model, its assumptions and requirements. At the time of writing, each protocol document spanned from 3 to 16 pages. A list of the protocols developed can be found in Appendix A.

2.4.2 Technical challenges

In developing the protocols, we encountered some challenges with the LLM platforms under evaluation. Perhaps most significant were issues related to the LLM context windows (often described as token counts or length), which is

the amount of information related to the current chat that the LLM can keep in ‘memory’ and that shapes its responses. These limits differ across platforms and vary across different versions, generally becoming larger with later versions. We needed to pre-load protocol documents into the LLM, and then required the LLM to retain this information and use it during the subsequent dialogue. This step necessarily uses up some of the available ‘memory’ of the LLM. As a result, we did experience issues that appeared to relate to a loss of information or context. Also, we often hit platform usage limits, which could mean the need to stop the dialogue (or pause for some timeout period). With Claude, hitting session limits could become limiting, however session lengths before hitting system limits seemed to be inconsistent. We speculated that other factors were in play (system demand management, wider system status issues). Where possible, we took steps to mitigate token usage, some of which used platform specific features. We found that the platforms would not always behave consistently week-to-week, or sometimes day-to-day, as model features and refinements would seem to be trialled by the platform provider on an experimental basis, even within ostensibly the same LLM version. Given the very fast development cycle of the LLM platforms, we had to accept this as part of the terrain (and acknowledge that some of these issues might now be anachronistic, while anticipating that new challenges might arise).

2.5 Data analysis

Two team members (*int.dev#1* and *int.dev#2*) analysed all session logs and chat records. To identify core ideas and themes regarding user evaluation and experience, they first conducted an overall holistic coding followed by an descriptive open coding of selected passages (Saldaña, 2016).

3 Results

3.1 Scaffolding modules targeting CT properties

We organised distinct scaffolding protocols into modules for the dimension of *Structural Complexity*, specifically for the properties (and corresponding sub-properties) of *Structural Variety and Dimensionality* (SVD) and *Relationality* (R). The sub-properties of SVD were all covered in the same protocol. For the property of *Relationality*, it was necessary to create two different protocols for the sub-properties of *Relations as Entities* and *Relational Movements*, as the processes of their scaffolding appealed to a different logic. The modules were constructed by composing and integrating different scaffolding modes. Figure 2a illustrates how the various modes are organised and related. Each specific

module would have its own sequence of modes, coordinated through its prompt protocol (see Sect. 3.2.2).

3.2 Scaffolding modes

3.2.1 Core scaffolding modes

Two modes of scaffolding assumed a central role in the organisation of each module: *Dialogical scaffolding* and *Evaluation*.

Dialogical scaffolding. The *Dialogical scaffolding* mode is the default mode of interaction with the user, building on the capabilities of the LLMs for natural language interactions. It is based on posing questions to the user and adjusting them to the level of complexity of their responses (according to the evaluations of the evaluation mode). We prepared a prompt protocol to support the user by presenting questions that would guide the thinking. The questions were selected and adjusted taking into account the indicators and levels of complexity for the property resulting from evaluation.

We prepared a bank of scaffolding questions to guide the LLM in developing and increasing the complexity of the users’ thinking. These questions were organised into different categories to be used in different moments of the scaffolding process.

For each target dimension, property and sub-property, we designed different categories of questions to be used in this mode, fulfilling different purposes, namely:

Particular categories of questions were created. For example, a set of modulating questions is embedded in the dialogical scaffolding that explores the user’s intentions with their thinking. Other questions modulate the thinking with other critical CT movements such as those associated with the dimension of *Complexity of the Observer*, including the properties of *Multi-Positioning*, *Reflexivity* and *Intentionalities*.

- (i) *Procedural questions*: to provide support to the user regarding the scaffolding process, e.g. offering instructions or a tutorial on the basic terms used;
- (ii) *Foundational questions*: to support the thinking in generating information for a minimally complex description of their TSoI;
- (iii) *Modulating foundational questions*: to specifically modulate the description of the TSoI through other properties, particularly regarding the *Complexity of the Observer*, including the properties of *Multi-Positioning*, *Reflexivity* and *Intentionalities*;
- (iv) *Scaffolding questions*: to increase the complexity of the thinking of the observer in terms of the target property;
- (v) *Modulating scaffolding questions*: to further increase the complexity of the observer, by incorporating other

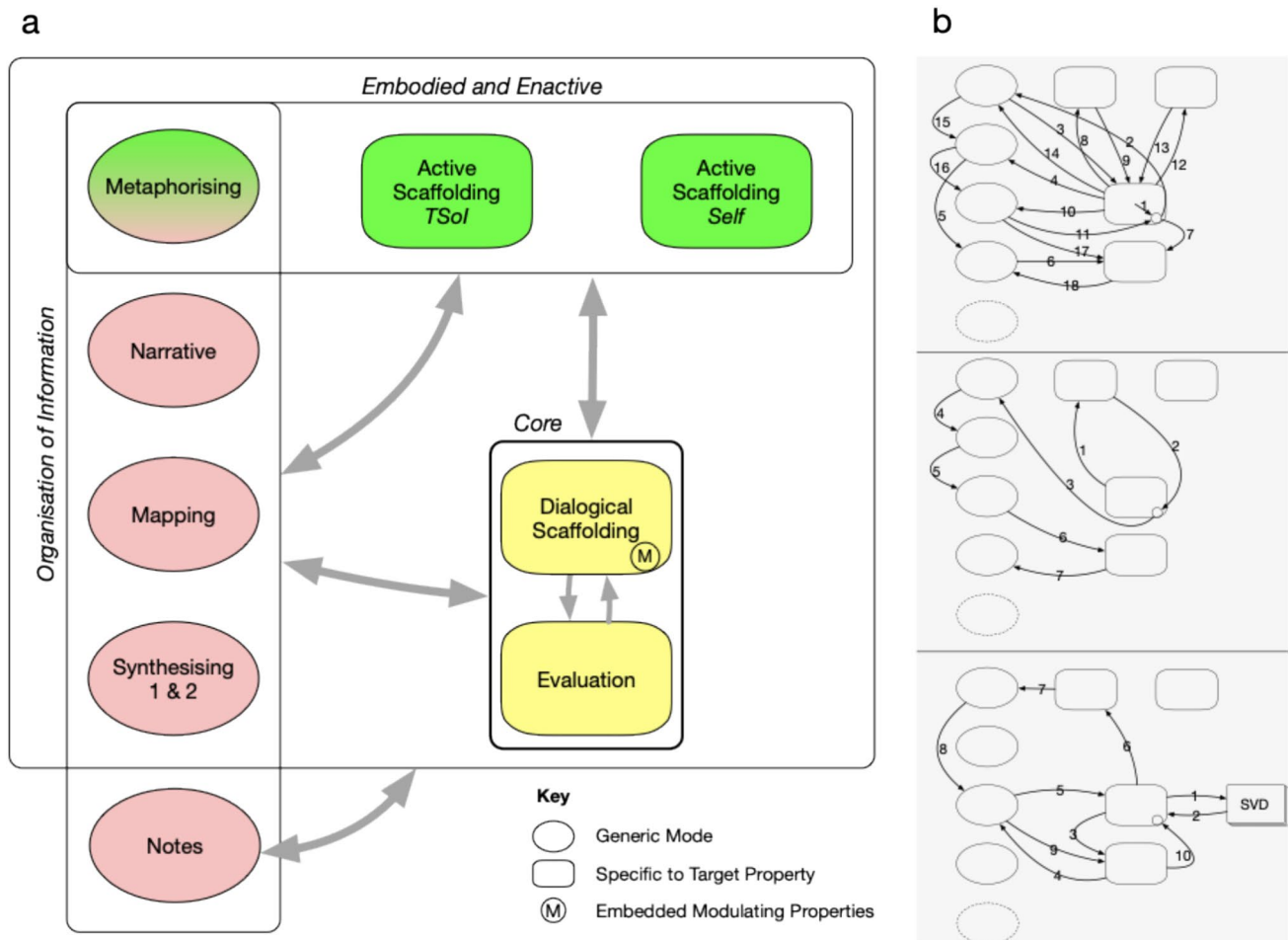


Fig. 2 Scaffolding modes: **a** Schematic organization of the 9 scaffolding modes that compose a scaffolding module targeting a CT property indicating their functional categories. The relations between the different categories of modes (e.g. interactions, transitions) are indicated with arrows, those within categories (excepting the Core) are not shown for clarity; **b** Illustration of sequences of movements through the scaffolding modes derived from the protocols for the CT properties of *Structural Variety and Dimensionality* (SVD) (top), and for the

sub-properties of *Relationality* (*R*): *Relations as Entities* (middle) and *Relational Movements* (bottom). For *Relational Movements*, the protocol incorporates outputs from prior explorations of SVD, as shown on the figure. The trajectories are indicative only, as the protocols also contain instructions to branch or recurse depending on the current evaluation, and also to make movements based on user choice. The sequence of moves is numbered

critical CT properties to modulate the target property. These questions were specific to each of the target dimensions and their corresponding properties and sub-properties.

Evaluation. The *Evaluation* mode provides a structured evaluation of the complexity of the expression(s) of the thinking of the user, according to the evaluation framework for CT for that particular property. It can be explicitly invoked by the user(s) to provide feedback. It is also implicitly invoked in the protocols. The *Evaluation* mode considers the outputs of all other modes in the evaluation of the complexity of the thinking expressed at a given point in the dialogical process. The evaluation results in an indication of the current level of the relative complexity of the

users' thinking, which affects all other modes by adjusting their responses (e.g. proposed activities). Evaluation is key to the scaffolding process by supporting attunement to the users' ZPD.

3.2.2 Embodied and enactive modes

The dialogical scaffolding was complemented with activities that connected the user with the embodied and enactive foundations of CT. These activities should create opportunities for a more direct engagement with non-verbal and non-propositional information. We created three modes of scaffolding: an *Active Scaffolding* mode focused on the *TSol*, an *Active Scaffolding* mode focused on *Self*, and a *Metaphorising* mode.

Metaphorising. Aims to support the user(s) in building a more embodied experience and holistic connection with the information, increasing reflexivity and facilitating the direct manipulation of the information generated. Generates syntheses of complex, nuanced and multidimensional information in the form of metaphors. Claude offers, unprompted, an exploration and unpacking of the metaphor, making explicit key aspects in relation to a complex conceptualisation of the TSoI. *Metaphorising* has a dual character in that it also functions in the organisation of information (see Sect. 3.2.3).

Active Scaffolding mode focused on the TSoI. Aims to support a direct engagement of the user with their TSoI (and its environments) in order to generate new meaningful information. It facilitates the creation and exploration of more embodied information generated in active direct engagement with the TSoI. The tool provides suggestions for different activities to be performed.

Active Scaffolding mode focused on the Self. Aims to support a self-referential exploration of the thinking, proposing active and embodied activities that allow the user to consider and explore their role as an observer in the construction of the TSoI and the nature of their own thinking. It promotes reflexivity and an embodied and situated stance.

3.2.3 Organisation of information modes

The CT framework sets requirements regarding the need to organise, relate, visualise and manipulate information. A set of modes was designed to provide a variety of ways of managing information, including a *Mapping* mode, a *Synthesising* mode, with two sub-modalities (cf. Table 3), a *Narrative* mode and a *Notes* mode.

Mapping. Contributes to increased reflexivity and awareness about the nature of the contents of the thinking considering the organisation of the TSoI. It highlights contents that are more and less developed and where absences are more critical.

Synthesising. Helps the user capture, organise and integrate the information creating visual and memory aids. It supports the management of complex information and aids the user in performing recursive visits and checks on the information generated. It has distinct sub-modalities that result in different artifacts.

Narrative. Helps the user capture the thinking developed so far in a narrative form, creating conditions for recursion and revisitation. It increases reflexivity and supports the sharing and coordination of the thinking with others external to the process.

Notes. Supports the user in the organisation of their thinking and helps increase reflexivity about its emergent outcomes. It facilitates the capturing, retrieval and integration of the moment by moment outcomes of the thinking, and the emergent information, into the overall process. It can

be used, at any moment, to capture information that can be further explored through different modes. The tool prompts the user to use the Notes mode but the user can also directly access it on request.

3.2.4 Mode artifacts

The complex and multi-layered information generated through these processes can be captured and synthesised in a variety of artifacts, generated by the *Organisation of Information* modes of *Metaphoring*, *Synthesising*, *Mapping* and *Narrative* which provide support to the scaffolding process through visual, diagrammatic, audio, and (movement of) narrative means. The resulting maps, syntheses and narratives can be subject to new evaluations (via the *Evaluation* mode) to provide multi-modal feedback to the users, which in turn enriches the basis of evaluation. The various Artifacts and their evaluation are reported in Sect. 3.4.

3.3 Configuration of scaffolding modules and sequencing of modes

3.3.1 Module configuration

Scaffolding modules for a target CT property include all the available modes in particular configurations of relations. Figure 2a shows the general organization of a prototype module with the 9 scaffolding modes. Of these core processes, *Dialogical scaffolding* mode is the most specific, in that it uses draws from a bank of questions that relate to the target property, but also includes modulating questions that embed other critical properties (e.g. those related to dimensions of the *Complexity of the Observer*) to modulate the target property. The *Dialogical scaffolding* mode is complemented by two distinct *Active Scaffolding* modes focused on the *TSoI* and the *Self*, that also utilise questions that are specific to the target property. The other modes are generic, and (we anticipate) can be applied to any of the CT properties.

3.3.2 Sequencing and management of scaffolding modes

The tool is composed of scaffolding modules targeting particular CT properties (or sub-properties). Each module contains prompt protocols that guide the user through a particular sequence of scaffolding modes. The protocols usually start with the *Dialogical* and *Evaluation* modes, which take a core role, which are then interspersed with the other modes. The protocols are composed of sequences of steps and recursive movements through which some modes are revisited, several times with different focuses and, in the case of the *Dialogical Scaffolding* with different categories

Table 3 Summary of scaffolding mode outputs/artifacts with potential improvements

Name and Identifier	Corresponding artifact	Potential improvements
Core modes		
<i>Evaluation</i> [EVAL MODE]	[EVAL REPORT] Generates a report identifying and justifying the level of complexity for each property and sub-property and an overall evaluation	Not identified
<i>Scaffolding</i> [DSCAFF MODE]	Not applicable	Provide the user with a visual mapping of the point where they are in the scaffolding protocol and the freedom to move to particular points. Allow an experienced user to skip the procedural and first foundational questions and to choose the types of scaffolding questions. Enrich the database of questions
Organisation of Information modes		
<i>Synthesising</i> [SYNTH MODE_OPTION 1] and [SYNTH MODE_OPTION 2]	Option 1 generates [TSOI_TABLE], a table organising the information about the TSOI according to the criteria of the theoretical framework for the contents that are necessary for a minimally complex description of the TSOI Option 2 generates [SYNTH MAP] a visual representation or mnemonic that helps the user grasp, capture and remember the essence of the information generated about the TSOI	Table layout need to be made more attractive Map could have different template layouts. Consider generating both Maps and Mnemonics, offering the user the option
<i>Mapping</i> [MAP MODE]	[TSOI_MAP] organises the information about the TSOI by the dimensions the framework considers necessary for a minimally complex description of the TSOI	Mapping of relations needs to provide more details and nuance. Map should also support the visualisation and outcomes of relational movements
<i>Narrative</i> [NARRATIVE MODE]	[TSOI_Narrative] Creates a narrative about the TSOI, capturing the thinking developed so far	Different kinds of narratives afford different possibilities. Options for style can be given to the user, e.g. literary narratives vs. journalistic / report
<i>Notes</i> [NOTES MODE]	[OBS NOTES] Captures the emergent ideas or insights the user has during the scaffolding process	Emergent ideas should be mapped onto the relational maps like the TSOI_MAP
Embodied and Enactive modes		
<i>Metaphorising^a</i> [METAPHOR MODE]	[METAPHOR MAP] AI-generated visual, audio, or video artifacts that capture the metaphors created by the user	Explore the potential of different categories of metaphors. Create dialogical scaffolding questions around the exploration of the metaphor
<i>Active Scaffolding TSOI</i> [ACTIVE SCAFF TSOI MODE]	Not applicable	A richer database of activities that support the enactment of a wide array of property and sub-properties of CT can be built drawing from different domains
<i>Active Scaffolding Self</i> [ACTIVE SCAFF SELF MODE]	Not applicable	A richer database of activities that support the enactment of a wide array of property and sub-properties of CT can be built drawing from different domains

^aMetaphorising is also considered an Organisation of Information mode

of questions (see Sect. 3.2.1). Example sequences of modes for the different modules are (for mode codes see Table 3):

Structural Variety and Dimensionality. This module integrates the following generic sequence of modes: DSCAFF Procedural questions > DSCAFF Foundational modulating questions > METAPHOR MODE > DSCAFF Foundational questions > NARRATIVE MODE > SYNTH MODE > EVAL MODE > DSCAFF Foundational questions > ACTIVE SCAFF TSoI or Self > DSCAFF Scaffolding questions > MAP MODE > DSCAFF Modulating question > ACTIVE SCAFF TSoI or (depending on user's choice) Self- DSCAFF Scaffolding questions > METAPHOR MODE > NARRATIVE MODE > MAP MODE > EVAL MODE > SYNTH MODE > User option to re-enter DSCAFF mode or move to another property.

Relationality. For this property it was found that different protocols were needed for the sub-properties:

Relations as Entities. This protocol integrates the following generic sequence of modes: DSCAFF Procedural questions > DSCAFF Foundational questions > ACTIVE SCAFF TSoI or DSCAFF Foundational Questions > DSCAFF Foundational modulating questions > METAPHOR MODE > NARRATIVE MODE > MAP MODE > EVAL MODE > User option to re-enter and continue DSCAFF mode or move to another property.

Relational Movements. This protocol follows this generic sequence of modes: DSCAFF Procedural questions > User upload of artifacts resulting from the scaffolding for *Structural Variety and Dimensionality*, including the TSoI_MAP, TSoI_NARRATIVE and/or TSoI_TABLE or to produce a narrative about their TSoI > DSCAFF Procedural questions > EVAL mode for Relations as Entities > MAP Mode > EVAL mode for Relational Movements > DSCAFF Procedural questions > DSCAFF Foundational questions > User option for ACTIVE SCAFF mode or DSCAFF Foundational questions > METAPHOR mode > MAP mode > EVAL mode > User option to re-enter and continue DSCAFF or end.

A simplified illustration of the trajectories through the modes can be seen in Figure 2b. Some of the relations and transitions between these modes are driven by the protocol but they can also be invoked on demand by the user, who is given, at the onset of the interaction, an overview of the different modes. Thus users can drive the dialogical interaction in particular directions. This freedom opens up the risk of the tool losing the immediate context of the protocol and/or straying outside of the guidance provided by the framework. Therefore, we added an instruction in the form of the prompt: "If the observer starts making other questions and requests that are unrelated or that move the dialogue away from the scaffolding protocol, and the different modes of scaffolding that are previewed, answer always in ways that would support the "Observer" [OBS] in increasing the

complexity of their thinking as described in the documents CT_CODING_STRUCTURAL_VARIETY_EVALUATION_MODE_NARRATIVE_INSTRUCTIONS, particularly considering the dimensions and sub-dimensions for a minimally complex description of the TSoI. Immediately after, resume the scaffolding at the last step that was performed according to instructions."

3.4 Scaffolding mode artifacts: evaluation and issues for improvement

Some scaffolding modes involve the generation of particular artifacts. Table 3 presents the 9 scaffolding modes with their corresponding artifacts, when applicable, and a brief description of aspects to be further developed or improved in the future based on the limitations that were identified.

As an example of attunement, the tool would make some natural adjustments in the artifacts in relation to the user(s) style of interaction. However, the final evaluation tests pointed to the relevance of making additional refinements to the *Narrative* mode to allow the user to choose the style of the narrative to be created (e.g. choosing between a more literary or more journalistic/succinct style). We found that different styles prompt/sustain different types of thinking and that they also can be used for different purposes. For example, the family support team of practitioners found the narratives to be very useful. While the literary style narratives offered them interesting ways to share their hypothesis with the families they worked in an effective and "beautiful" or inspiring way, the more journalistic narrative style helped them structure information in ways that were useful in preparing reports (e.g. assessment case reports to send to courts or child protection systems).

3.5 Pilot testing results: Strengths, limitations, and suggested improvements

Each case study had a different context and focus, with the exception of 5 and 6, which had similar foci. In Appendix B we describe the context and purpose of each case study, highlighting the most distinctive positive and negative features. In Appendix C we present selected excerpts of the chats, illustrating particular modes of scaffolding and some of the most salient features of the interaction of the user with the tool. Overall, the case studies contributed to further refinements and adjustments of the protocols.

The analysis of all the data (videos, transcripts, and chat logs from the session, as well as logs with memos made by observers of the sessions) allowed us to identify some patterns regarding the more salient features of the tool (LLM + protocol suite) and their effects. Table 4 summarises these patterns in terms of the strengths and limitations of the tool (Claude Sonnet 3.5 + protocols) and their effects,

Table 4 Perception of the Tool: Strengths and limitations (Claude Sonnet 3.5 + protocol suite)

Strength	Magnitude	Limitation	Magnitude
Interaction			
[the interaction with the LLM feels like] “Natural”, fluid engagement	High	[the user experiences the interaction as] “Verbose”	Low
Good coupling: Tool acts as reflexive partner staying close to and supporting the user’s thinking but not thinking for them, while introducing some difference and variation	High	Technical information [e.g. identifiers] is not omitted	Low
Good adaptation: adapting questions and activities to context	High	Coupling can be lost in the transition to other protocols when procedural questions and foundational questions are not revisited to create a stronger connection with the TSoI	Moderate
Transition between protocols facilitated by uploading artifacts from previous sessions	Moderate	Transition between protocols is poor (Properties)	High
Returns to protocol after some deviation from it (e.g. addressing user question) making the connection with the closest scaffolding questions	Moderate	May not return to protocol promptly or when “pushed” significantly out of it	Low
Element of surprise and stimulation of curiosity that supports engagement	High	Sometimes presents too much information having an “overwhelming” effect [on the user]	Moderate
Mostly consistent performance with chose model	Moderate	Some erratic, “off order” but “in protocol” mode interactions	Low
		Model-specific performance of the protocol	High
Questions: <i>Dialogical Scaffolding</i>			
Generative, productive, interesting questions	High	Sometimes [the responses of the LLM are perceived as] “massive” [by the user] and difficult to answer (immediately)	Moderate
Adaptation and choice of questions to user context and preferences	High	Calls upon a concrete context of application or TSOI	Low
Generates engagement, sustains attention	Moderate		
Relationality associated with new perspectives and insights, and creative effects	Moderate		
Scaffolding through metaphors: <i>Metaphorising</i>			
Metaphors support the tool-user coupling, keeping the dialogue close to the user frame of reference	High	Sometimes offers metaphors spontaneously	Low
Metaphors situate the thinking and support the connection with rich embodied information; Brings the user closer to their own thinking and their enhances their capability of interacting with it	High		
Tool “unpacks”, explores the metaphor in a ways that add and highlight other or implicit perspectives and nuances	High		
Metaphor supports a richer coupling of the tool/-user and user with their TSoI	High		
Updating of the metaphors supports connection the thinking and its updating	High		
Visualisation of information: <i>Mapping</i> and <i>Synthesis</i> modes			
Good, useful and rich mnemonics	High	Often requires additional prompting to offer mnemonics (protocol needs refinement)	Moderate
		Quality and detail of the MAPs of the TSoI variable	Low
Good structured syntheses, following evaluation framework	High	Poor visual layout of syntheses tables and other synthetic artifacts	Low
Ampliative effect of the syntheses (adds nuances and generates new information through increasing reflexivity)	Moderate	Syntheses are sometimes trivial or considered unnecessary when information is limited and they are “just recaps” [the user does not identify much novelty]	Low

Table 4 (continued)

Strength	Magnitude	Limitation	Magnitude
Organising and reflecting back: <i>Narrative</i>			
Offers rich narratives which increase reflexivity and can have multiple uses outside of the interaction	High	The style of the narrative seems to be influenced by the style of expression of the user. Would be useful to have narrative modes on demand (e.g. journalistic vs. literary)	Low
Narratives integrate and elaborate on metaphors expanding the thinking. Ampliative effect of the narratives	Moderate		
Scaffolding effects			
Effects on the thinking: Inducing reflection, adding new perspectives, organising, stimulating new modes of thinking (e.g. loops); Thought provocation that may continue beyond the session	High	Requires time (to think and to respond) and may require several interactions at multiple times	Moderate
Stimulates and offers support for new interactions and modes of engagement with the TSol	High	May generate frustration when only limited amounts of information can be processed (e.g. vs. uploading massive amounts of information)	High
The user knows the modes available and can take an active role by calling for them and for artifacts	Moderate	The user is not aware of “where they are” in the (overall steps of the) protocols so it is difficult to choose pathways	Low

classifying our perception of their magnitude as strong, moderate and weak.

Following the sessions for cases 5 and 6, we asked the tool to provide a detailed overview of the steps of protocol that were (or were not) followed. It provided a thorough response and justification of where the protocol was adhered to and where there were variations, justifying the adaptations appropriately given the guidance for scaffolding CT and the need to adjust to the team’s cases and nature of the thinking.

In most of the tests, the users engaged with the supporting members of the team, not just asking for technical advice on how to proceed but also to share their thoughts out loud or reflect upon them. As we did not have a control condition, it is not possible to explore the role and effects of this human presence and their impact on the scaffolding process. In cases 5 and 6 there was a discussion between members of the team which was triggered by the questions of the tool. The effects of this discussion are likely to be relevant but the absence of comparative conditions does not allow us to elaborate on them.

In all cases the exploration of the *Relationality* protocols was more limited. For this property, the tool has not been able to generate interesting artifacts (i.e. visual representations that are informative and navigable) that could be used in a mapping mode to record the relational movements performed and the relations identified in the conceptual space that describes the TSol.

The users and the observers of the tests provided direct and indirect suggestions for improvement and further innovations regarding the protocols and use of the tools. Some of these correspond to adjustments that can be made to current protocols, likely with the result of reducing effort and difficulty such as: (i) allowing the user more choices in regarding the type of syntheses artifacts to be created and the style of the narratives; (ii) allowing the user to choose a style of interaction with the tool (e.g. fact-oriented, synthetic and “less polite” interaction vs. friendly, verbose interaction); (iii) providing the user with a map of the scaffolding process allowing them to be aware of which stage they are at and which steps they can “jump” to. This last suggestion poses some dangers as it is necessary to ensure that the user does not skip any stage that is necessary to build the foundations for performing CT. Others suggestions may imply more technical challenges and other kinds of user interfaces. For example, one user suggested having more texture and levels of interaction i.e. having the option to see all the information but also to choose which level to explore in more depth. He also suggested a “learning diary” or journal, where the user could keep a daily record of their thinking stimulated by the tool, with this additional information being used to enrich the interactions. Other suggestions, e.g. allowing the tool to analyse and respond to large amounts of information, are constrained by technical difficulties and the capabilities of

the models such as the constraints posed by context memory capacity, or file upload limits, or the capacity of the tool or indeed the user(s) to handle very large datasets (e.g. on a target system). Finally, we believe it is relevant to develop protocols for both novice and advanced users, guiding them in the use of the tool in a responsible manner, while warning of the limitations of the current tools for scaffolding CT within the CT framework..

4 Discussion

4.1 Challenges in scaffolding CT: Epistemic authority

There is a tendency to consider AI as capable of generating thinking independently, holding knowledge more advanced than humans, and answering our questions and solving our problems. In other words, there is a trend towards endowing AI with epistemic authority and decisional sovereignty (Ferrario et al. 2024; Bartsch et al. 2024). Our main concerns, and the challenges we faced in this project, were related, directly or indirectly, to the preservation of a coupling between humans and the AI that enhances the thinking of the human, while safeguarding the user from what we consider a potential epistemic harm in the delegation of their decisions to AI (Malone et al 2024; Fricker 2007). In designing our tool, we sought to address these challenges: the initial scaffolding questions in the protocol specifically target the observer's context, clarifying their intentionalities and grounding the thinking in them. Throughout the whole process, we aimed to ensure that decision-making remained in the user's hands and that any interpretive suggestion by the tool was understood as such — merely a suggestion and a possible interpretation or introduction of a perspective — which could and should be assessed by the user for its pertinence and relevance. This requirement seems to have been met, as evidenced in the testing (e.g. users' comments on case 4).

4.2 On the development and evaluation of protocols to scaffold CT

We identified the requirements and challenges for scaffolding CT in ways congruent with the conceptual foundations of the CT framework. We then undertook a phased process of development of the prompt protocols designed to scaffold a selection of key properties of CT (covering all properties was not feasible in the timeline of the project). The most successful AI models in addressing complex problems use a combination of approaches and integrate different kinds of models (Marcus & Davis 2021). We used a combination of modes of engagement and scaffolding to make the tool

more capable of supporting and enhancing the type of thinking that could address complex problems. For the protocols, we developed a way of organising scaffolding modules (for particular CT properties) by integrating a variety of different scaffolding modes. We identified key elements of our approach, which led to a scaffolding mode, which in turn relied upon the construction of a robust evaluation mode that could guide the behaviour of the tool and the user. We found other modes that supported this process, in supporting the organisation of information generated and promoting the embodied and enactive forms of knowing.

We tested the protocols on different LLMs, using our guiding criteria: seeking to stay within the boundaries of the CT framework, achieving attunement to the user and context, and introducing variation. In many ways, ChatGPT was the most advanced platform, having certain technical advantages, such as better graphical output, a larger context window and more generous usage limits. However, we kept coming back to Claude. With Claude, we achieved the best balance of our guiding criteria using our protocols and it was also the most engaging dialogical partner.

It is interesting to note that Anthropic, the developers of Claude, prioritise maturity and safety: they incorporate a list of rules or principles during the AI training stages, resulting in what they call Constitutional AI (Bai et al 2022). Moreover, Anthropic employs a philosopher to oversee Claude's character training (<https://www.anthropic.com/research/claude-character>). Interestingly, Claude is reported to be a favourite amongst people within the tech field, citing its insightful nature and “emotional intelligence” (Roose 2024). Additionally, the use of Artifacts (persistent, editable “documents”) afforded by Claude was critical to the success of the tool as they provide a way for the user to actively interact with the outputs of their own thinking in a timely and intuitive way.

Despite the current limitations of the tool, the users' experiences point towards effects that were desirable from the perspective of the CT Framework. The detailed step by step algorithmic approach to the guidance and instructions sections of our protocols provided the LLM with a sort of “learner model” and strategy for an ITS (Sætra 2022); the general first instructions, evaluation protocols, introduction, and appendixes can be regarded as comprising a “teaching module” (Sætra 2022) on CT, with the LLM providing the dialogical interface.

Some of our prompts share characteristics with patterns previously identified as effective by other authors, such as White et al. (2023), e.g.: reflection patterns (ask the model to explain the rationale of its answers); context manager (specify or remove a context for the conversation); recipe pattern (providing constraints to output a sequence of steps); visualisation generator pattern (use text to create visualisations); template pattern (provide a precise template for

structuring an output); cognitive verifier patterns (subdivision of questions into additional questions); and the alternative approaches pattern (ensure LLM offers alternatives to user).

In general, the final protocols led to desired effects and outcomes on the targeted properties. They suggest the viability of developing systems of coupled AI and human intelligences where AIs are “partners in cognition” (Salomon et al. 1991) that are capable of (at least partially) performing some key scaffolding functions in relation to the performance of CT, such as motivation, direction maintenance, simplification, marking critical features and demonstrating (Sætra 2022).

4.3 On the possibilities of interaction between a CT framework and AI (LLM) tools

We aimed to explore the possibilities of interaction of a CT framework and AI, specifically LLMs, and to develop prompt protocols to scaffold CT. The experience of operationalising CT in the context of this study, provided insight into the role and interplay of the different dimensions and properties of the CT framework. Some dimensions adopt a more central role (structural, dynamical, causal/explanatory and dialogic complexity), in particular the structural dimension, which generates the *contents* of the thinking. Others are seen as modulating properties (observer, developmental/adaptive, pragmatic, ethical/aesthetical, and narrative complexity) which are movements that *shape and fine tune* the thinking, helping to increase its fitness to the specific conditions being considered. These insights will be incorporated into a revision of the CT framework (Melo & Renault, in preparation). Hence, it is not only necessary to develop protocols for other properties of CT but also meta-protocols (and associated tools) to coordinate their interaction and coordination and to choreograph their performance.

We designed the protocols considering, in particular, the need to scaffold the thinking of a particular observer with a TSoI to afford more complex understandings which could, eventually guide more effective decision-makings and interventions. The tests seemed to provide richer results with users that had a more invested relationship with their TSoI and a concrete context of application/practice where an intervention could take place. At present, we consider that the tool may work better in case-based scenarios, that are more concrete, than for more abstracted levels of thinking, likely as a consequence of grounding the ideas in their particularities, making them more tangible and manipulable. The results provide some support to the idea of distinguishing between the effects of thinking with AI and the effects of AI (Salomon et al. 1991); user’s reports suggest transformations in patterns of thinking that may be triggered in

direct contact with the AI and yet continue to be performed afterward.

The results are encouraging enough to suggest extending the approach to the other properties of the CT framework. However, some challenges remain: for example, the need for better mapping and visualisation of complex information (e.g. multiple levels, multiple degrees of depth and relations) associated with the thinking movements, particularly in support of *Relational Movements*. Our experience, for example, in the Complexigraphy method (Melo & Renault 2023), shows us the importance of this kind of visual feedback to the user(s) in becoming aware of their patterns of thinking. This current limitation of the tool, we believe is a result of (the then current) technical capabilities of the LLMs, and can be overcome as the platforms become more mature and better integrated with other tools (e.g. network/data visualisations). Moreover, with more resources, a more customised tool could be developed using the LLM API.

Any system to scaffold CT necessarily needs to be open and to support rich and close modes of direct and indirect coupling between the user, their TSoI and their environments. This includes the alignment and ecosystemic fitness of the thinking. This can be done through key modulating dimensions (e.g. ethical, aesthetic or pragmatic complexity) which are dependent on evaluations and inputs provided by critical observers in such systems. The construction and valuing of meaning related to the thinking also needs to be retained on the side of the user and other observers in their TSoI with whom dialogues need to be established.

4.4 The users' experiences through the scaffolding of CT

Our results point to a general positive experience of the users in the interaction with the tool, which they identified as valuable. In general there was an experience of a good coupling between them and the tool. We need to explore the possibility of having different modes of interaction and types of interfaces allowing for different types of engagement over time between the user and a tool. For example, designs that support the prompting, production, management and integration of the different types of interactions on an “everyday basis” to support and enrich more complex direct couplings of a user with their TSoI and to support an even more dynamic and ongoing unfolding and steering of the thinking. The protocols may benefit from adjustments to create versions for first time and novice users and for more experienced and advanced users. There is also the challenge of designing customised protocols for particular types of Systems and users and to assess differences and similarities of critical processes in different contexts. Our tests suggest the need to further explore and investigate different configurations of hybrid systems of co-intelligence. For

example, the role of a Human Scaffolder agent alongside the AI scaffolding role. There is a potential positive role for the mere presence of another human, e.g. to create a favourable emotional environment, increase the confidence and engagement of novice users, as well as to increase reflexivity and to amplify the thinking processes stimulated by the tool. The effects of the use of these tools in teams and the differences between individual and collective uses needs investigation.

5 Limitations

In this short exploratory project, only a small number of exploratory case studies were conducted. This poses several limitations on the interpretation of our results and on our conclusions, particularly those concerning objectives related to evaluating users' experiences. Further studies should conduct more pilot evaluation studies with a larger and more diverse sample of participants, including participants with various degrees of experience with LLMs, familiarity with CT, and different intentions in relation to their TSoI. More systematic studies are necessary to test the robustness of the protocols under different conditions of application (e.g., with and without human support) and to evaluate outcomes in terms of the Complexity of Thinking achieved and the impact on users' follow-up actions. Experimental and quasi-experimental designs will be warranted to explore variations resulting from different conditions of protocol application and with different groups.

We only conducted a broad evaluation of general aspects of the UX. Future studies should consider a more systematic exploration of a wider number of dimensions including issues related to trust, credibility of the results, as well as studies on the impact of the systematic use of these kinds of tools. We developed protocols for scaffolding two individual properties of CT. However, the framework highlights the centrality of the interaction between properties. Through introducing modulating questions in the scaffolding modes we were able to find a proxy for working for these interactions. However, we have not been able to address how to have the LLM manage and guide the transitions between modules/properties and how to integrate them. This is a significant limitation that warrants deeper investigation.

6 Conclusion

We explored the possibilities and constraints of the interaction of a framework for the practice of CT with AI, specifically LLMs, through the development of a tool to scaffold particular CT properties of observer in relation to a TSoI. Our results suggest the viability of using such tools, under well defined constraints set up by a robust evaluation

framework for CT and a detailed, narrative and algorithmic prompt protocol. The suite of prompt protocols developed encompass a variety of modes of scaffolding aiming to achieve a balance of keeping within the bounds of the CT framework, while adapting to the user's context, and adding new information. We found the performance and user experience of such a tool is dependent on the choice of LLM. Many technical challenges remain to be overcome, in particular the development of meta-tools to support the coordination and integration, orchestration or choreographing (Melo & Renault 2023) of CT. Some challenges may be naturally overcome due to developments in the LLMs. Others may require the combination and integration of a variety of technologies and types of AI, including protocols for the assemblage of systems with different configurations of co-augmented intelligences (Melo 2022), e.g. incorporating (i) multiple users/observers (e.g. teams), attending and facilitating the complexity of their own coupling and (ii) the coordination of other human and AI scaffolders. The possibilities are vast, and as systems for scaffolding CT evolve, so may our understanding of scaffolding and CT performance itself. We believe we have taken a small but important step towards the realisation of tools to scaffold CT. This will support the development of more complex modes of thinking and action in addressing the wide variety of complex challenges that we face, globally, locally, individually and collectively. Amongst these, we believe that the scaffolding of CT could play a role in the training of AIs, for example in addressing the complex problem of the alignment of their goals and behaviours with different systems of human values and intentions.

Appendix A. List of protocols

A protocol suite was developed to Scaffold CT in terms of the Dimension of Structural Complexity, and its properties of Structural Variety and Dimensionality and Relationality. The following protocols were integrated and fine-tuned for Claude Sonnet 3.5 (information between squared brackets refers to the identifiers associated with the protocols):

1. GENERAL FIRST INSTRUCTIONS [SC.SVD.REL. GENERAL FIRST INSTRUCTIONS_EVAL AND SCAFFOLDING] protocol, to be inserted into projects (containing different versions of the text, depending which of the other protocols are used) (3 p.);
2. Evaluation of Structural Complexity/Structural Variety and Dimensionality [SC.SVD. EVALUATING STRUCTURAL VARIETY AND DIMENSIONALITY] (5 p.);
3. Scaffolding Structural Complexity/Structural Variety and Dimensionality [SC.SVD. SCAFFOLDING STRU

- CTURAL VARIETY AND DIMENSIONALITY] (16p.);
4. Evaluation of Structural Complexity/Relationality [SC.REL_EVAL_ EVALUATING RELATIONALITY] (8p.);
 5. Scaffolding Structural Complexity/Relationality/Relations as Entities [SC.REL. R. SCAFFOLDING RELATIONS] (8p.);
 6. Scaffolding Structural Complexity/Relationality/Relational Movements [SC.REL.RM. SCAFFOLDING RELATIONAL MOVEMENTS] (8p.).
 7. Integrated Evaluation and Scaffolding of Relationality, Relations and Entities and Relational Movements [SC.REL.RE.RM.EVALUATING AND SCAFFOLDING RELATIONALITY] (14 p.).

Appendix B. Context, purpose and highlights of case studies

Case 1: int.user#1

Context/purpose. The user had a theoretical-applied research problem. He was experienced with AI. He operated at the intersection of academic research and the justice and mental health system by analysing and consulting on the decisions related to the execution of security measures by mentally disordered offenders. He was aiming to have insights on the subject. There were several test sessions, continuing the same chat.

Negative highlights of the interaction. Outputs were sometimes considered “verbose”.

Positive Highlights of the interaction. i. Experience of a “natural” dialogue; ii. “Enjoyable experience”; iii. First quite good experience with radial maps (improved after this test); iv. Interesting mnemonics were created; v. Notes mode was very active. Tool identified insights in the user and proposed to use [NOTES MODE]; vi. Return to Relationality after a session break requires a “warm-up” and procedural questions stage; vii. Interaction, particularly Relationality triggered new insights and perspectives; viii. Presence of the team possibly made the non-linearity of the dialogue more manageable and enjoyable than with an individual exploration; ix. Syntheses were considered useful (“It’s very, very interesting to see this whole information synthesized in a single page. It’s very helpful”).

Case 2: int.user#2

Context/purpose. The user works in the domain of social psychology. She was inexperienced with AI. She had a theoretical research problem focused on housing issues. She was

interested in understanding how to enrich her research from a CT perspective. There were several test sessions, opening new chats on each occasion.

Negative highlights of the interaction. i. The user was “impatient” with the tool. The user did not wish to answer nor write down detailed responses to questions or to repeat “information she already knew” but wanted feedback on her own work and suggestions for improvement, resulting in sometimes frustrating experience; ii. User wished to upload large amounts of information (e.g. book written on the issue) and have feedback from that but it rapidly used up token limits for the day and shortened the interaction; iii. User deviated frequently from the protocol avoiding the direct questions of the tool; iv. User found having to write responses “tiring” and would have preferred oral interaction.

Positive Highlights of the interaction. i. Questions considered interesting, stimulating new ideas; ii. Sometimes “exciting” experience.

Case 3: ext.user#1

Context/purpose. The user was experienced with AI. He is a researcher interested in Futures research. He defined his topic and system of interest as “uncertainties of the future of work”.

Negative highlights of the interaction. i. Outputs were sometimes considered “verbose”; ii. User would prefer identifiers to be hidden; iii. Mapping felt as a bit “too obvious” but still helpful (“Documenting, not really adding anything. (...) But it helps. OK”); iv. Problem was too abstract. Could benefit from concrete examples (which was suggested by the tool); v. Sometimes questions perceived as “too massive” i.e. difficult to respond to immediately.

Positive Highlights of the interaction. i. Metaphor exploration was well-received and useful, interpretations are close to user meanings but offer new perspectives (“yes, very good, catalyst of change [tool offered interpretation of the metaphor] I like that (...) yes, this is nicely mapping out what I’ve been thinking in different ways, this is good”); ii. Positive experience of novelty and insights introduced through the Relationality/Relational Movements (“for me this part is more interesting, to get new perspectives and new starting points and then to connect them”; “Actually, it’s interesting because it’s inviting you to think about them the other way around”); iii. Appreciation of how tool “keeps track of everything and brings them together”. There is an indication of a more comprehensive and inclusive view of the problem, which integrates various elements, thereby facilitating not only relational movements but also a more visual organization of information. There is an appreciation of its ability to maintain an overview of complex information.; iv. Appreciation of the questions (“the questions

are quite good”; “yes, I think these are good statements”); v. Questions were well adapted to context (“it’s just doing that on its own. This is perfect then”); vi. Appreciation of embodied focus (“yeah, yeah, it’s good”); vii. Experience of recursive thinking loops (“It’s making me think of a kind of a loop now”); viii. Team experience that they were the ones doing the thinking, and that their curiosity was stimulated, and they were engaged, supported by the tool (“yes, that [it] makes us think. It is indeed very interesting, very complementary and challenging, but at the time it is structuring (...) truly provokes in us, I want to know more, let me see now what you are going to say, how we can do this [implement the suggested activities] What can I do with the family?”).

Case 4: ext.user#2

Context/purpose. User was inexperienced with AI. She is a researcher working with socio-ecological problems. She defined her problem as “how the sociopolitical dimension of nature-based solutions can be explored using complex thinking tools”.

Negative highlights of the interaction. i. User had expectations that her questions would be answered; ii. Glitch: too many questions presented at once, although corrected later on; iii. Some modes were skipped. User’s discourse somehow prompted to go to “Relationality” too early.

Positive Highlights of the interaction. i. User had insights of new ideas (“great to think about”); ii. User appreciated suggestions of activities to carry on. iii. User engaged in a critical evaluation of the questions and didn’t take the information for granted (“I think this question, it’s better than the other, because there are some of these elements more influential than others in shaping the best outcomes. I think this is a great point to think about it”). There is a sense of thought provocation that does not appear to conclude with the interaction with the protocol but promises to continue over an extended period.

Case 5: ext.user#3

Context/purpose. A team of two users of community-based family support practitioners conducted the test focused on a case conceptualisation of a family with at-risk children. The team had two separate sessions, the second on Relationality.

Negative highlights of the interaction. i. The transition between protocols in the two different sessions was perceived as “cold”, lacking the “warming-up” and focus on building a connection with the case of the first session on the Structural Variety and Dimensionality protocol; ii. The

tables layout was poor and uninteresting (“but it is a shame it was not in a [proper] table format”).

Positive Highlights of the interaction. i. The unpacking and exploration conducted by the tool of the metaphor proposed by the team was considered as “highly potent” and “an excellent resource to use with the family”, likewise the narratives (in literary-style) that were produced based on the metaphor (“it can be a very useful intervention tool with the family (...) I believe this will have a very positive impact on her”); ii. Synthesis of information, including in the tables was considered useful despite the poor layout (“the information is well synthesised in topics, which guides the organisation”). The way the synthesis is organised and feeds back information adds something new (but close) to the thinking of the team (“I thought the same magic happened a moment ago, this and that. I really liked the synthesis (...) It is a way for us, new to an immediate situation, to receive a very well-organised synthesis [when sometime the thinking still isn’t] which we then use for information; it is true, but we receive the information organised, so it is, wow!”, “besides the bullet point referencing the tool created a relating by areas, which organises the information and guides the way we are thinking the case, amplifying our perspective”, “gives us a perspective of what is missing”); iii. “Natural” interaction and engaging coupling and “game-like” interaction (“it happens in an almost natural way, doesn’t it? We became involved, didn’t we? So now it’s a question [posed by the tool], now I do it this way. And it was something very fluid, wasn’t it? Something that almost then emerged in a very natural way. And I didn’t feel the weight of the technology”).

Note: Although we developed the protocols in English, the users interacted with the tool in Portuguese. The tool dealt with the translations. The results were positive.

Case 6: ext.user#4

Context/purpose. A team of two users of community-based family support practitioners conducted the test focused on a case conceptualisation of a family with at-risk children. The team had two separate sessions, the second on Relationality.

Negative highlights of the interaction. i. Same as case 5.

Positive Highlights of the interaction. i. The team appreciated and used the [NOTES MODE] frequently; ii. The questions posed by the tool triggered discussions between the team members generating a variety of complementary perspectives. The team presented these diverse perspectives back to the tool for exploration; iii. The team frequently used the team member supporting the test as another scaffolding and reflexive partner and discussing their choices with her; iv. The metaphors and reports were highly valued, enriching the thinking (“it was very good, it was very enriching. The spectacular metaphors, the way it...(did it)”);

iv. Productive and enriching experience; v. The way the tool fills in and adds what is not explicitly said (but was implicit in the text and in the metaphor chosen) (“*I found it to be very productive. Thank you,, thank you as well. Thank you, we understand so many aspects. Thank you, goodness. We write and don't write everything. And the way it manages to read what we are writing, I found spectacular. And believing that what we did not say ended up being reflected in the resolutions that the program was making.*”).

Note: The users’ natural language is Portuguese. They interacted with the tool in English using a translator app on their browser. Although we developed the protocols in English, the users interacted with the tool in Portuguese. The results were positive.

Appendix C. Illustrative excerpts from chat logs for the case studies

Case 1 excerpts: int.user#1

Early stage of the dialogue introducing procedural questions

Participant	Excerpt of the chat log
AI/ Claude	Excellent. Let's begin with Phase 0, which aims to clarify procedures and language [PHASE 0] [STEP 0.1] I'll use the [PROCEDURAL QUESTIONS] from Appendix A to confirm and clarify the working conditions for our scaffolding session PQ1. Would you like to have a small introduction tutorial on key terms that we could use to structure our conversation? PQ2. I propose to call your problem/concern/question "the Target System of Interest (TSoI)". If you would like to call it something different, please tell me Could you please respond to these procedural questions so we can establish a clear foundation for our scaffolding process?
User	My relation is as a consultant/advisory to a court
AI/ Claude	Thank you for sharing that information about your relation to the TSoI. Let's continue with the next foundational modulating question to further explore your perspective [SC.SVD.MD.O.MOD.OC.MP.FMODQ2] What are your intentions in relation to your TSoI? What would you like to achieve? For example, would you like to build (i) a comprehensive description, (ii) explanations, (iii) anticipations of predictions (e.g. of what could happen if nothing or something changes or possible effects of interventions)?
User	I am trying to make a decision on whether an individual diagnosed with a mental disorder, currently compensated, should be eligible for parole
AI/ Claude	Thank you for clarifying your intention. This helps us understand the context and complexity of your TSoI. Let's continue with the next question to further explore your perspective [SC.SVD.MD.O.MOD.OC.MP.FMODQ3] What is important or interesting to you in relation to your TSoI?
User	It is important to know: 1. if the individual will maintain treatment; 2. if the community is prepared and willing to accept him
AI/ Claude	Thank you for sharing these important aspects of your TSoI. Your focus on treatment maintenance and community preparedness shows a consideration for both the individual and the broader social context. Let's move to the final foundational modulating question to complete our initial exploration [SC.SVD.MD.O.MOD.OC.MP.FMODQ4] What are the values guiding your interaction and intentions in relation to your TSoI?
User	My values consider the overall common good. That is, to regard the individual's rights (to liberty and autonomy), while protecting the community concerns for safety

Early stage of the dialogue introducing procedural questions

Participant Excerpt of the chat log

Asking for the first metaphor [METAPHOR MODE]

AI/ Claude Thank you for articulating the values guiding your interaction with the TSoI. Your focus on balancing individual rights with community safety demonstrates a nuanced approach to this complex situation

Now that we've completed the foundational modulating questions, I'd like to help you think of a metaphor that best represents the nature of your relation with the TSoI

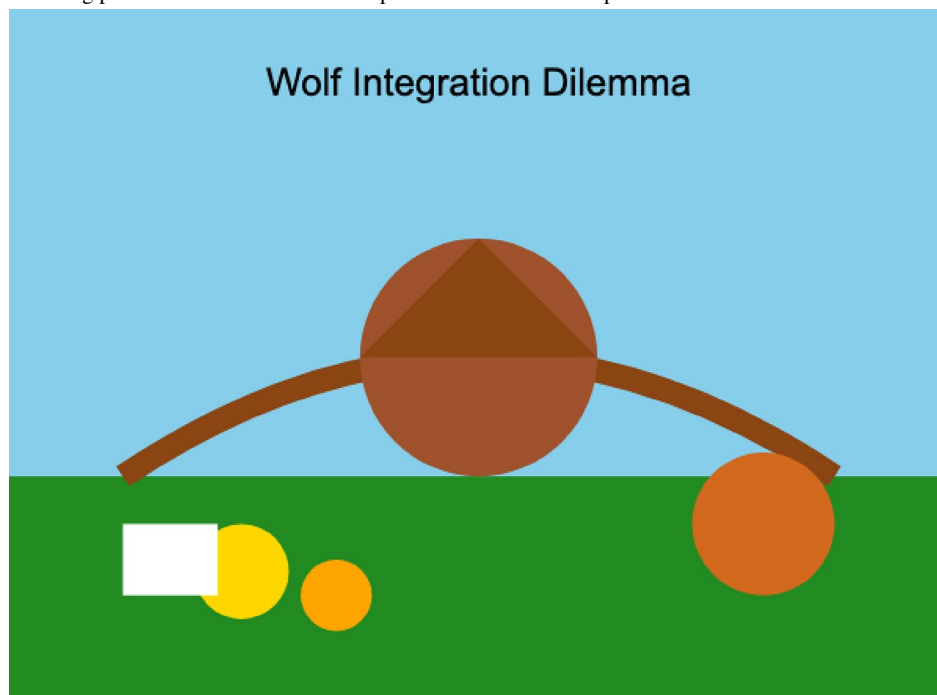
[METAPHOR MODE]

Can you think of a metaphor (e.g., an image, a word, a sound, an object, or a movement) that captures the nature of how you are thinking about this decision-making process regarding the individual's eligibility for parole?

User So, you have a wolf that has been trained and is being adopted by a family that has other animals, like dogs and cats, but also livestock. Can we trust the training, or that the wolf will come out, eventually?

Unpacking and exploring the metaphor on [METAPHOR MODE]

AI/ Claude Thank you for sharing this powerful metaphor. It effectively captures the complexity and potential risks involved in your decision-making process. Let's create a visual representation of this metaphor



This visual representation captures the key elements of your metaphor:

1. The wolf (representing the individual with a mental disorder) is shown as larger and separate from the other animals, symbolizing its potential threat and the uncertainty of its integration
2. The family is represented by a house, with various animals (dogs, cats, and livestock) nearby, symbolizing the community
3. A fence separates the wolf from the family and other animals, representing the current separation and the potential for integration
4. The sky and grass background represent the broader environment in which this situation is taking place

This metaphor effectively illustrates several key aspects of your TSoI:

1. The trained wolf represents the individual diagnosed with a mental disorder who is currently compensated
2. The training of the wolf mirrors the treatment the individual has received
3. The family adopting the wolf represents the community that would be accepting the individual back
4. The other animals (dogs, cats, livestock) represent different members of the community with varying levels of vulnerability
5. The question of trusting the training or the potential for the wolf's nature to "come out eventually" directly relates to your concerns about whether the individual will maintain treatment and whether the community is prepared for potential risks

This metaphor encapsulates the balance you're trying to strike between the individual's rights (the wolf's chance at a new life) and community safety (protecting the family and other animals). It also highlights the uncertainty and potential risks involved in your decision

Now that we have this metaphor to work with, let's move on to [PHASE 2] to further explore different dimensions of the complexity of your thinking about this TSoI

Early stage of the dialogue introducing procedural questions

Participant Excerpt of the chat log

Initiating foundational questions in Dialogical Scaffolding mode [DSCAFF]

AI/ [PHASE 2]
 Claude [STEP 2.1]
 I'll now guide you through a series of [FOUNDATIONAL QUESTIONS] to help you reflect on different dimensions of the complexity of your thinking. We'll start with the first question:
 [SC.SVD.MD.IC.FQ1] Please provide an initial description of your Target System of Interest (TSoI), considering its basic elements, such as types of individuals and groups as well as other types of beings, existing hierarchies and levels of organization among them, nature of their interactions and relations, different functions and purposes supported by these elements, patterns of the TSoI and how they relate back to the elements, narratives and preferred identities of the TSoI
[Because of a glitch the AI did not confirm the metaphor before proceeding. The user calls for improvements]
 User Can you use objects instead of geometric shapes?
 AI/ Of course! I apologize for the abstract representation. Let's create a more concrete visual representation of your metaphor using
 Claude objects instead of geometric shapes
 (...) **[RESUMING FOUNDATIONAL QUESTIONS AFTER IMPROVING METAPHOR REPRESENTATION]**
 AI/ Now, let's return to our foundational question. Could you please provide an initial description of your Target System of Interest
 Claude (TSoI), considering its basic elements, individuals, groups, their interactions, and the overall patterns and narratives involved in this parole decision scenario?

Illustration of scaffolding questions

Participant The individual is a male, under 30, suffering from schizophrenia, that received a security measure because he murdered his brother
 AI/Claude **[CONTINUING SCAFFOLDING WITH FOUNDATIONAL QUESTIONS]**
 Thank you for providing this specific information about the individual in your Target System of Interest (TSoI). This gives us a clearer picture of the complexity involved. Let's continue to build on this description with the next foundational question:
 [SC.SVD.MD.E.FQ2] Please provide a description of the social, material, cultural, institutional and natural environment of TSoI
 Could you describe the environment in which this decision is being made? This could include aspects such as:
 ● The institutional context (e.g., the court system, mental health facilities)
 ● The social and cultural context of the community where he might be released
 ● Any relevant laws or policies that govern this decision
 ● The available support systems or resources for individuals with mental health conditions in the community

Reminding the user on the possibility of using [NOTES MODE]

AI/Claude Also, I'd like to remind you that during our interaction and the scaffolding process, you're encouraged to pay attention to ideas that come to your mind and their potential relevance for understanding your TSoI. Try to "think with your whole body" and pay attention to your reactions, emotions, or even vague impressions. At any moment, you can ask me to make a note of these ideas and impressions by activating the [NOTES MODE]. Just let me know if you want to do this at any point
 Are you ready to proceed with the scaffolding process?

Scaffolding Relationality/Relational as Entities

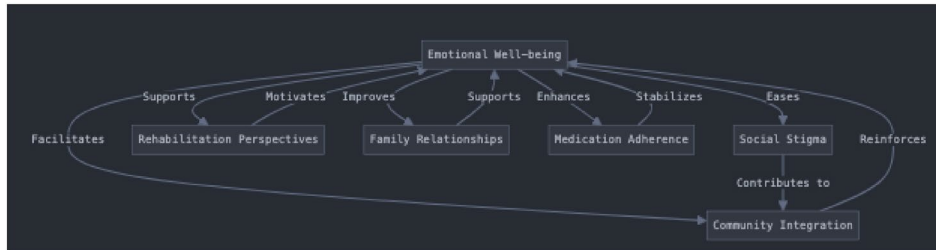
AI/Claude Now, let's consider how these aspects relate to other elements in our TSoI:
 [SC.RL.R.F.Q3] How can the relations between the key elements in your Target System of Interest (TSoI) be described? What other types of relational descriptions, qualifiers and types, dimensions or properties of relations can be considered?
 Considering the emotional well-being aspect you've brought up, how would you describe its relationship to other key elements we've discussed (such as medication management, family support, community perceptions)? Are there any new types of relations or properties that come to mind when considering emotional well-being?
 User Well, emotional well-being could support relations with the environment, namely, by easing the social stigma and integration into community groups. In the long term, I believe it would benefit rehabilitation perspectives

Scaffolding Relationality/Relational as Entities

AI/Claude Excellent observation. You've identified an important relationship between emotional well-being and several other key elements of the TSoI. Let's explore this further:

[SC.RL.R.F.Q4] Are there any patterns or structures that emerge from these relations?

Your insight suggests a potential pattern: emotional well-being as a facilitator for other positive outcomes. This could form a structure where emotional well-being is at the center, influencing various other aspects of the TSoI's life and rehabilitation. Let's visualize this:



This diagram illustrates the central role of emotional well-being and its bidirectional relationships with other key elements of the TSoI

Now, let's delve deeper into these relationships:

[SC.RL.R.F.Q5] How do different types of relations interact and influence each other? Are there relations between relations?

How do changes in one relation affect other relations within the TSoI?

For example, how might improvements in the TSoI's emotional well-being affect his relationship with his family, and how might that, in turn, impact his medication adherence or community integration? Can you identify any potential chain reactions or feedback loops in these relationships?

Offering a mnemonic [SYNTH MODE] after an evaluation of Relationality [Relations as Entities and Relational Movements] followed by an updated map of the thinking [TSol MAP]

User Can you give an easy to memorise version of this?

AI/Claude Certainly. I'll provide a concise, easy-to-memorize version of the relationality evaluation using a simple framework [ARTIFACT]

TIES Relationality Summary

T—Threads: Multiple interconnected relationships

- Illness insight and emotional well-being form the central thread
- Many threads connect various aspects (family, treatment, community)

I—Integration: Holistic view of the TSol's situation

- Different elements are woven together cohesively
- Changes in one area ripple through the entire system

E—Exploration: Dynamic investigation of relationships

- Circular thinking examines how factors influence each other
- Considers both immediate and long-term effects of relationships

S—Sophistication: High complexity in relational thinking

- Qualifies relationships (e.g., "deeply intertwined")
- Identifies relationships between relationships

Remember: The narrative TIES together multiple threads, integrates various elements, explores dynamic relationships, and shows sophistication in relational thinking

This TIES framework provides an easy-to-remember summary of the key aspects of relationality in your updated narrative:

1. Threads remind you of the multiple interconnected relationships you've identified
2. Integration highlights how you've woven different elements together into a cohesive whole
3. Exploration emphasizes the dynamic nature of your relational thinking
4. Sophistication underscores the high level of complexity in your consideration of relationships

By remembering TIES, you can quickly recall the main strengths of your relational thinking about the TSol. This framework also provides a quick checklist for maintaining and further developing the complexity of your relational thinking in future analyses

User Can you give me an updated map of my thinking?

Offering a mnemonic [SYNTH MODE] after an evaluation of Relationality [Relations as Entities and Relational Movements] followed by an updated map of the thinking [TSol MAP]

AI/Claude Certainly. I'll create an updated map of your thinking based on our recent discussions, particularly incorporating the central role of illness insight and emotional well-being, as well as the various relationships we've explored [ARTIFACT: Updated TSol Thinking Map]



This updated map of your thinking about the TSol reflects the increased complexity and relationality we've discussed. Here are the key features:

1. Central TSol: The young male with schizophrenia remains at the center
2. Core Focus Area: Illness Insight and Emotional Well-being are now highlighted as a core focus area, represented by the dashed ellipse encompassing the central TSol
3. Four Main Dimensions:
 - 1.A Internal Complexity (right)
 - 1.B Environment (left)
 - 1.C Observer (top)
 - 1.D Coupling (bottom)
4. Sub-elements: Each main dimension has relevant sub-elements, reflecting the depth of consideration for each aspect
5. Relationships: Arrows indicate the relationships between different elements, with a focus on how the core area of Illness Insight and Emotional Well-being connects to other dimensions

This map visualizes several key aspects of your thinking:

- The centrality of illness insight and emotional well-being in your understanding of the TSol's situation
- The interconnectedness of various elements, showing how changes in one area might affect others
- The balance between internal factors (symptom management, medication adherence) and external factors (family dynamics, community integration)
- The role of the observer (court consultant) in assessing and potentially influencing the system
- The coupling between treatment effectiveness and rehabilitation progress

This visualization captures the TIES framework we discussed earlier:

- **Threads:** Multiple interconnected relationships are shown by the arrows
- **Integration:** The central focus area integrates with all other dimensions
- **Exploration:** The layout suggests dynamic relationships between elements
- **Sophistication:** The inclusion of sub-elements and multiple relationships demonstrates the complexity of your thinking

This map serves as a visual summary of your current thinking about the TSol, highlighting the complex, relational nature of your analysis. Indicates how these aspects of the observer's role are related

Illustration of a response to a user's direct request for a solution from within the framework and returning to the dialogical scaffolding mode [DSCAFF], offering suggestions as possibilities alongside questions

User	Let me take you aside. How would you organize a system to care for not criminally responsible patients, in the described context?
AI/Claude	Thank you for this thought-provoking question. To address this, let's consider a systemic approach that takes into account the complex relationships we've discussed. I'll outline a potential organizational structure, and then we can explore how it might address the issues we've identified [SC.REL.RM.SCAFF.Q18] To what extent could these relationships be integrated or explained by a larger or higher-order entity or dimensions or be related at the level of their parts? Consider this potential organizational system: 1. Integrated Assessment Centers: - Multidisciplinary teams including legal experts, psychiatrists, psychologists, and social workers - Conduct comprehensive evaluations to determine both legal status and medical needs 2. Specialized Treatment Facilities: - Separate units for different types of conditions (e.g., one for schizophrenia, another for brain injuries) - Staff trained in both legal implications and medical treatment of specific conditions 3. Tiered Insight Development Program: - Tailored programs to develop illness insight based on cognitive capabilities and specific conditions - Regular reassessments to adjust legal status and treatment plans 4. Integrated Legal-Medical Review Boards: - Regular case reviews by a board of both legal and medical professionals - Ensures ongoing alignment between legal status and medical treatment 5. Transitional Care Units: - Bridge the gap between secure facilities and community reintegration - Focus on practical application of illness insight and management 6. Continuing Education Program: - For staff across all units to stay updated on both legal and medical developments - Promotes a culture of integrated legal-medical understanding Now, let's explore how this system might address the relationships we've discussed: 1. How might this integrated approach affect the development of illness insight in patients with different conditions? 2. In what ways could this system better balance legal requirements with medical needs? 3. How might this structure influence the relationship between healthcare providers' approaches and patients' understanding of their conditions? 4. How could this system potentially resolve some of the tensions we identified between legal categories and medical treatment approaches? 5. In what ways might this organizational structure itself emerge as a new factor in the relationship between illness insight and the healthcare system?

Case 2 excerpts: int.user#2

Illustration of visual synthesis map after the production of a narrative

AI/Claude

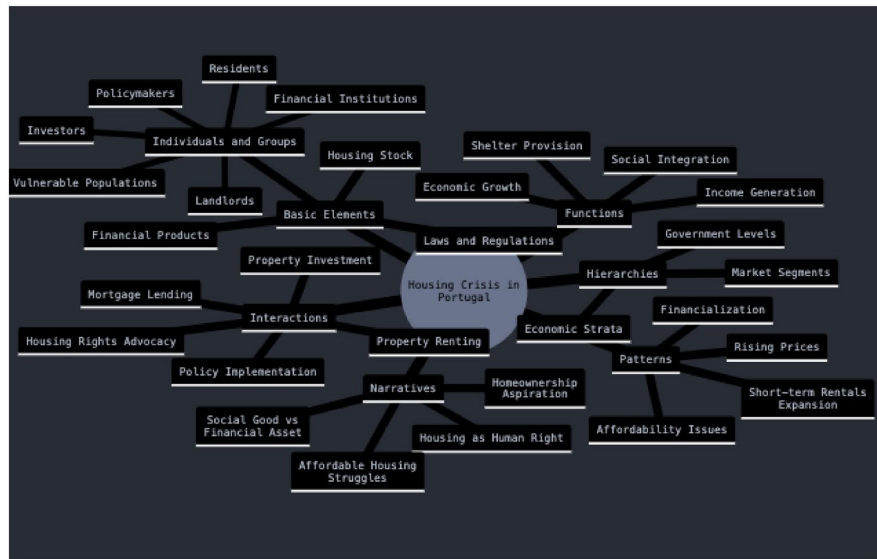
Excellent. I'm glad the narrative accurately reflects your understanding of the housing crisis in Portugal. Let's move forward with our analysis

[STEP 2.3]

Now, I'd like to offer you a synthesis of the information we've generated so far

Synthesis: Housing Crisis in Portugal

[ARTIFACT]



I've created a mind map to synthesize the key elements of the housing crisis in Portugal. This visual representation aims to capture the main components and their relationships within the system

The mind map is organized around the central concept of the "Housing Crisis in Portugal" and branches out into the main categories we've discussed: Basic Elements, Hierarchies, Interactions, Functions, Patterns, and Narratives. Each of these branches further subdivides into more specific elements

This synthesis can help in visualizing the complexity of the system and the interrelations between its various components.

It can also serve as a quick reference for the key aspects of the housing crisis

Would you like me to explain any part of this synthesis in more detail? Or shall we move on to the next phase of our analysis?

User

Thank you. Let's continue

Certainly, let's continue with our analysis and move on to the next phase

Illustration of a follow-up explanation of the evaluation provided

AI/Claude (following presenting an evaluation and a TSoI_TABLE) Would you like me to explain any part of this evaluation in more detail?

User Yes

AI/Claude Certainly. I'll explain each part of the evaluation in more detail:

[TSoI_TABLE]: This table maps the contents of your thinking onto four key dimensions: a) Internal Complexity of TSoI (1.A): This is the most developed dimension in your narrative. It includes a rich description of basic elements (like residents, landlords, investors), hierarchies (government levels, economic strata), interactions (mortgage lending, property renting), functions (shelter provision, income generation), patterns (financialization, rising prices), and narratives (homeownership aspiration, housing as a human right). b) Environment of TSoI (1.B): This dimension is less developed. While you implicitly reference economic conditions and cultural values, there's limited explicit description of broader environmental factors that influence the housing crisis. c) Observer (1.C): This dimension includes your role as a social psychology researcher, your intentions to build comprehensive descriptions and explanations, and your values of fairness, inclusivity, and sustainability. d) Coupling (1.D): This dimension is the least developed, with limited explicit description of how the system, environment, and observer interact

Sub-properties Evaluation: a) Multidimensionality and Variety of Contents [MDV]: Rated as Moderate complexity. Your narrative covers three out of four dimensions (internal complexity, observer, and some environment) with high thematic variety in the internal complexity. However, the environment and coupling dimensions are less developed. b) Depth of Information: Rated as Moderate complexity. You provide various types of information (structural, functional, narrative), but the sources of information are not explicitly varied. c) Nature of Contents: Rated as Moderate complexity. Your narrative balances individual and relational entities, and atomic and configurational descriptions. However, it leans more towards abstract descriptions with fewer concrete examples. d) Framing the Information: Rated as Moderate complexity. The information is somewhat contextualized (e.g., mentioning changes since the 1990s, urban vs. rural differences) and adds meaning to understanding the TSoI. However, more detailed spatial and temporal contextualization could enhance this aspect

Overall Evaluation: Your thinking demonstrates moderate complexity. It excels in describing the internal complexity of the system with a rich variety of elements, interactions, and patterns. To increase complexity, you could:

- Explicitly address the broader environment of the system (e.g., global economic trends, climate change impacts on housing)
- Describe the coupling between system components, the environment, and yourself as an observer
- Provide more concrete, situated examples to balance the abstract descriptions
- Explicitly mention diverse sources of information

Is there any specific part of this evaluation you'd like me to elaborate on further?

Illustration of suggestion of activities [ACTIVE SCAFFOLDING] for theoretical problem of housing

AI/Claude [STEP 3.4]

Given that we've explored the environment dimension to some extent, I'd like to offer you the opportunity to engage in some active scaffolding to potentially increase the complexity of your thinking and produce novel information on your Target System of Interest (TSoI) Would you like some suggestions for activities that could help you further develop your understanding of the housing crisis in Portugal, particularly focusing on the environmental aspects we've been discussing?

These activities could help you generate new insights or connections, especially in areas that are less developed in your current thinking, such as the broader global context and its interaction with local factors

If you're interested, I can suggest a few activities. If not, we can move on to the next phase of our analysis. What would you prefer?

User Yes, Please suggest a few activities

Illustration of suggestion of activities [ACTIVE SCAFFOLDING] for theoretical problem of housing

AI/Claude	Certainly! I'm glad you're open to some activities to further develop your understanding of the housing crisis in Portugal, particularly focusing on the environmental aspects. Here are a few suggestions: Active Scaffolding Activities: Housing Crisis in Portugal [ARTIFACT] [ACTIVE TSoI MODE] Active Scaffolding Activities: Housing Crisis in Portugal 1. Global–Local Mapping Exercise Create a visual map that illustrates the connections between global financialization trends and local housing markets in Portugal ● Use a large sheet of paper or a digital mind-mapping tool ● In the centre, place "Portuguese Housing Crisis" ● On one side, list global factors (e.g., international investment trends, global economic policies) ● On the other side, list local factors (e.g., local housing policies, demographic changes) ● Draw lines connecting these factors to the central topic, using different colors or line styles to represent different types of influences (direct, indirect, strong, weak) ● Add brief notes explaining each connection This activity can help visualize the complex interplay between global and local factors, potentially revealing new connections or areas for further research 2. Stakeholder Perspective Analysis Conduct a role-playing exercise to explore different stakeholder perspectives on the housing crisis ● Identify 5–7 key stakeholders (e.g., international investor, local resident, policymaker, social housing advocate, real estate developer) ● For each stakeholder, write a brief (1–2 paragraph) statement from their perspective about the housing situation in Portugal ● Consider their motivations, concerns, and how they might view other stakeholders ● After writing these, compare and contrast the perspectives. Look for areas of conflict and potential common ground This activity can help develop a more nuanced understanding of the various interests at play in the housing crisis, potentially revealing new angles for policy approaches or research questions 3. Historical Timeline Construction Create a timeline of key events and policies related to housing in Portugal, with a focus on how they interact with global trends ● Start from a significant point in the past (e.g., Portugal's entry into the EU) up to the present ● Include both local events/policies and global economic milestones ● For each entry, briefly note its impact on housing affordability or accessibility ● Look for patterns or turning points where global and local factors intersect This activity can help identify historical trends and pivotal moments that have shaped the current housing crisis, potentially revealing long-term patterns or cycles that could inform future predictions or policy recommendations 4. Comparative Case Study Select 2–3 other countries or cities facing similar housing challenges and conduct a brief comparative analysis with Portugal ● Choose locations with some similarities to Portugal (e.g., in size, economic situation, or housing market structure) ● For each case, summarize key aspects: housing market trends, policy approaches, social impacts ● Create a table or matrix comparing these factors across the cases ● Identify unique aspects of Portugal's situation as well as common threads This activity can provide a broader context for Portugal's housing crisis, potentially revealing alternative approaches or unique factors in the Portuguese case that might not be apparent when studying it in isolation These activities are designed to engage with the environmental aspects of the housing crisis in Portugal, encouraging you to make new connections and generate novel insights. They address areas that were less developed in your initial description, such as global–local interactions and historical contexts ___ [END ARTIFACT] I've suggested four activities that could help you further develop your understanding of the housing crisis in Portugal, with a focus on environmental factors: Global–Local Mapping Exercise Stakeholder Perspective Analysis Historical Timeline Construction Comparative Case Study Each of these activities is designed to help you explore different aspects of the housing crisis, particularly focusing on the interaction between global and local factors, diverse stakeholder perspectives, historical context, and comparative analysis These activities aim to increase the complexity of your thinking by encouraging you to: Visualize complex relationships Consider multiple perspectives Analyze historical trends and their impacts Compare Portugal's situation with other contexts Would you like me to explain any of these activities in more detail? Or do you have a preference for which activity you'd like to focus on?
-----------	---

Illustration of Dialogical Scaffolding, selecting and adjusting questions to context

User	Thank you Claude. The main alternative approach is considering that the housing crisis is a crisis of lack of houses. Housing is unaffordable because the construction of houses is low and therefore the low offer leads to high prices
AI/Claude	Thank you for sharing this important alternative perspective. It's a significant contrast to the financialization-focused view you presented earlier. Let's explore this further [SC.SVD.MD.MOD.OC.MP.SCAFFMODQ6] Given this alternative perspective that the housing crisis is primarily a supply issue, who would you identify as the critical observers or entities implicated in the housing crisis in Portugal? Please identify all the key stakeholders, including those who might hold this alternative view of a supply shortage Consider including: ● Those who advocate for increased construction ● Those who focus on financialization ● Government officials at various levels ● Construction industry representatives ● Urban planners ● Economists with different viewpoints ● Affected residents (both homeowners and renters) ● Any other key groups you think are relevant How do these different groups contribute to or are affected by the housing crisis?

Case 3 excerpts: ext.user#1

Illustration of Narrative Mode integrating previous metaphor (trickster)

AI/Claude	Based on our dialogue, I'll create a narrative that captures how you're thinking about your TSoI. After I present it, please let me know if you'd like to make any amendments or if it accurately reflects your thinking so far # Narrative: Uncertainties of the Future of Work The exploration of the uncertainties in the future of work is a complex and multifaceted endeavor, characterized by the interplay of various actors, spaces, and evolving concepts of value and meaning. This Target System of Interest (TSoI) is approached from the perspective of a researcher who, while not an expert, is deeply engaged with the topic and seeks to facilitate meaningful discussions about possible futures Key elements of this TSoI include: 1. Actors: Workers, non-workers, and others, including potential influencers like policymakers and technology developers 2. Spaces of Work: A spectrum from neutral, flexible spaces to structured, potentially controlling environments. The concept of work spaces is being redefined by factors such as seasonal work, night work, and day work 3. Value and Meaning of Work: These are in constant flux, influenced by changing social norms and external factors like environmental and political influences 4. Environmental Factors: Climate change is increasingly shaping the future of work, potentially leading to radical shifts such as nocturnal work schedules in hotter climates 5. Cultural and Social Responses: Adaptability and resilience are key themes, with examples like Cuba's social response to power outages highlighting the potential for reimagining work-life integration 6. Institutional and Material Aspects: There's a noted tension between slow institutional adaptation and the need for rapid change in response to environmental and social shifts The approach to this TSoI is characterized by: 1. A 'trickster' mentality that challenges established assumptions and introduces unpredictability 2. A focus on human skills, relationships, and agency in the face of technological changes 3. A critical stance towards current work structures, combined with willing participation in them 4. An emphasis on seeing alternatives and embracing the tension between dualities 5. A view of the unknown future as both a possibility and a threat, balanced with the recognition that current knowledge and experiences can guide engagement with uncertainty The ultimate aim is to explore scenarios and anticipate new realms to aid critical discussion on the current state of affairs, with a particular emphasis on valuing human aspects over technological fixes and maintaining an open future where people have agency This narrative summarizes our discussion about your TSoI, "uncertainties of the future of work." Does this accurately reflect your thinking so far? Would you like to make any amendments or additions to this narrative?
-----------	---

Illustration of procedural questions for scaffolding Relational Movements

User	Can you please help me by scaffolding relational movements
AI/Claude	Certainly! I'd be happy to help you scaffold relational movements to further enhance the complexity of your thinking about the uncertainties in the future of work, particularly focusing on the impact of climate change. We'll use the instructions from the CT_CODING_SC.REL.RM_SCAFFOLDING RELATIONAL MOVEMENTS PROTOCOL to guide this process Let's begin with [PHASE 0] to clarify procedures and language [SC.REL.RM.PQ.Q1] I would like to invite you to explore a wider number and type of potential relations between the bits of information and the different dimensions and sub-dimensions for the description of your TSoI. Are there relations that you particularly would like to explore and for which I could provide you with relational movements scaffolding questions? If so, we will start with those relations and later on I will propose we explore relations randomly. How would you like to proceed? [SC.REL.RM.PQ.Q2] I will call each bit of information or dimension that we will consider to explore in a relational way, a "relatum" or "relata" (plural). It means an information entity that has an inherent relational nature and is better understood in the context of other information or dimensions. Can we proceed? [SC.REL.RM.PQ.Q3] As I pose the questions, I will invite you to pay attention to what ideas come to your mind and to the potential relevance of the relations explored. Try to "think with your whole body" and pay attention to your reactions. If when making a relational movement another dimension comes to mind, or interesting or novel information (e.g., hypotheses) or ideas, let me know so that I can make a note. You just need to ask me to add a note to a mental map. I can also record those ideas as Relational Movement Ideas and maybe guide new relational movements from there. I will number and visually map the Relational Movement ideas in the Relational Map we are creating. Once you feel that you have a truly emergent idea (a new hypothesis, an insight) or something that comes to mind that you feel might be important and that is not simply reduced to the relata, I can add it to the map, mark it and number it as an Emergent Idea. That way we can later see the relations we built, with the emergent ideas and to the base ideas, to test its fitness. Can we proceed? Please let me know if you're comfortable with these procedures, and we can then move on to exploring relational movements

Illustration of a selection and adaptation of a Scaffolding Question for the property of Relationality/Relational Movements

AI/Claude	Thank you for that insightful response. You've highlighted the complex, circular nature of the relationship between climate change and work patterns, especially in agriculture. This circular perspective is particularly interesting. Let's explore this further: [SC.REL.RM.SCAFF.Q7] How does time determine or affect the nature of the relation between these relata? How are these relata related temporally (e.g., co-exist, one follows the other, have different rhythms, there is recursion, etc.)? Consider how the relationship between climate change impacts and work patterns/industries might evolve over different time scales. Think about immediate responses, medium-term adaptations, and long-term transformations. How might these temporal aspects influence each other?
-----------	--

Illustrative excerpts for Case 4: ext.user#2

Illustration of a selection and adaptation of a Scaffolding Question for the property of Structural Variety and Dimensionality

AI/Claude	Thank you for your thoughtful response, B. You've identified several important additional elements in the system of nature-based solutions (NBS). Let's build on this to deepen our understanding Based on the elements you've identified, let's explore how they're organized within the system: [SC.SVD.MD.IC.2.SCAFFQ2] How are these elements (political structures, economic factors, social effects, and cultural diversity) internally arranged in relation to each other in the context of NBS? Are there hierarchies or distinct levels of organization among them? Take your time to consider this question. Your response will help us understand the structural complexity of the sociopolitical dimension of NBS
-----------	---

Case 5 excerpts: ext.user#3

Illustration of a Dialogical Scaffolding followed by Synthesis

- AI/ Claude [SC.SVD.MD.E.FQ2] Could you provide a description of the social, material, cultural, institutional and natural environment of the 'love family'? For example, what is the neighbourhood like where they live, what community resources do they have available, or what environmental challenges do they face?
- User The housing conditions are quite precarious in structural terms, but the mother takes care of the space in a very brave and careful way, transforming it into a home. The context is somewhat challenging, in that it is mostly precarious and inhabited by families with low resources and other associated problems, which sometimes lead to conflicts. One of the family's main aims is to move out of the house, but financial conditions haven't allowed them to do so. They have already tried to find other solutions through social housing, but there has not yet been a positive response. Despite these conditions, the area has good sports and leisure facilities as well as easy access to the transport network and other services
- AI/ Claude Thank you for this detailed description of the Amor 'family' environment. This information adds important layers to our understanding of the system. I will synthesise and organise this new information:
1. Housing conditions:
 - Structurally precarious
 - Mum transforms the space into a home through meticulous care
 - Strong desire to move to better housing
 2. Social context:
 - Mostly precarious neighbourhood
 - Neighbourhood with low-income families
 - Occasional conflicts in the community
 3. Economic conditions:
 - Limited financial resources
 - Difficulty in moving to better housing
 - Active search for solutions through social housing (unsuccessful so far)
 4. Community infrastructure and resources:
 - Good sports and leisure facilities
 - Easy access to transport network
 - Proximity to other services
 5. Resilience and adaptation:
 - Mother's ability to create a cosy environment despite structural limitations
- This description of the environment significantly enriches our understanding of the 'Amor' family, showing how external factors influence its dynamics and well-being
- To continue developing a complex understanding, I'll ask the next fundamental question:
- [SC.SVD.MD.O.FQ3] Please provide a description of yourselves as observers who are thinking about and interacting with the 'love family'. Do you consider yourself to be part of, in contact with, or outside the 'love family'?

Illustration of Narrative Mode

- User please enter narrative mode and offer me a TSoI_NARRATIVE
(Note: The chat took place in Portuguese. All text was translated from Portuguese to English. Translated was performed with DeepL.com (free version))

Illustration of Narrative Mode

AI/Claude Certainly. I'll go into narrative mode and offer you a TSoI_NARRATIVE based on all the information you've shared so far (Note: the user defined its TSoI as a family case, they called Family Amor, which means Love)

[ARTIFACT]

TSoI_NARRATIVE: The Story of the Amor Family

In the heart of a challenging neighbourhood, where economic and social difficulties are a constant, we find the Amor Family. This single-parent family, made up of a resilient mum and her 11-year-old son, is like a whirlwind that moves love around, facing adverse winds, but always in search of light

The mother, marked by a history of parental rejection and an accident that left visible scars, carries deep insecurities. However, she has transformed these painful experiences into a fierce and protective love for her son. Her life revolves around him, creating a cosy home even in precarious housing conditions. She keeps everything impeccably clean, as if external order could compensate for the internal chaos that sometimes plagues her

Her son, a docile and polite boy, navigates between the role of child and confidant. Well integrated at school, he bears the brunt of his mother's worries, secretly wishing she could relax a little more. The recent change of school cycle has brought his mother's anxieties to the surface, threatening the autonomy and self-confidence the boy has been building

The father, although he doesn't live with them, is a constant presence. His more permissive parenting style contrasts with the mother's authoritarian approach, creating a complex dynamic. Both parents, however, share a deep love for their son, each expressing it in their own way

The family lives in a delicate balance, like an isolated microcosm in the midst of a larger community. Their social circle is small, limited to a few close family members and trusted friends. This reduced social network, together with limited access to information, creates a protective but also isolating bubble

In this scenario, the support team emerges as a vital bridge. You are more than observers; you are active participants in this family system. Your presence brings not only practical support, but also a connection with the outside world, mediating tensions and opening up possibilities

The Amor Family is a testimony to resilience and unconditional love. Each day is a delicate dance between protection and the need for growth, between the desire for stability and the inevitability of change. Like the whirling wind that symbolises their journey, they continue to spin, facing challenges but always seeking the light of love and hope

The future of the Amor Family remains uncertain, but full of potential

With the continued support and empathetic understanding you offer, there is hope that they can find a healthier balance. The challenge lies in helping the mother loosen the reins a little, allowing her son to grow in autonomy, while strengthening the support network around them

Ultimately, the story of Familia Amor is a powerful reminder of the complexity of human relationships and the resilience of the human spirit. It's a story still in development, with many chapters yet to be written, where every small step towards growth and stability is a victory to be celebrated

This TSoI_NARRATIVE narrative offers a holistic and contextualised view of the Amor Family, integrating the various elements and dynamics we've discussed so far. It captures the complexity of family relationships, the impact of the environment, and your role as a support team

Some key points that emerge from this narrative:

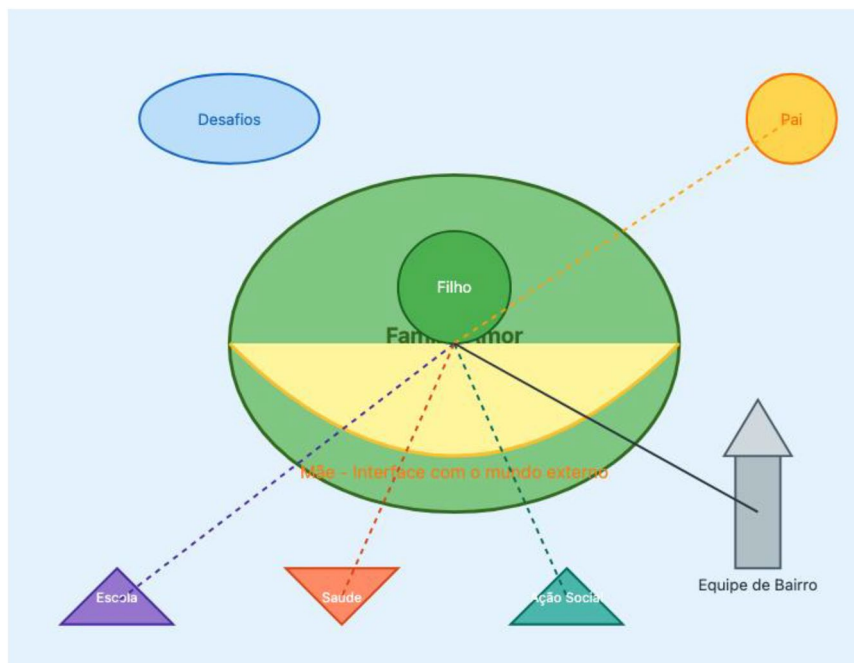
1. The family's resilience in the face of multiple challenges
2. The complex dynamic between protection and the need for autonomy
3. The impact of social and informational isolation
4. The importance of your role as a bridge between the family and the outside world
5. The potential for growth and positive change, despite the difficulties

This narrative can serve as a reference point for future reflection and to guide your interventions. It can also be updated as new information and developments emerge

Would you like to make any changes or additions to this narrative? Or would you prefer us to continue exploring different perspectives on the family, as suggested in the previous question?

Illustration of the MAPPING Mode combined with METAPHOR MODE

User	Can you please make an updated version of the TSoI map?
AI/ Claude	Of course. I'm going to create an updated version of the TSoI (Target System of Interest) map for the 'Amor' family, incorporating the new insights and the metaphor of the paradise island that we developed [ARTIFACT]



This updated TSoI map integrates the paradise island metaphor with the factual elements of the 'love family' case. Here is an explanation of the elements:

1. island (centre): represents the 'love family' as a whole
2. Beach (edge of island): Symbolises the mother as the interface between the family and the outside world. The wavy shape represents the protective and sometimes turbulent nature of this interface
3. Vegetation (centre of the island): Represents the son, well adapted but vulnerable to changes in the environment
4. Sun: Symbolises the father, a constant but distant presence, influencing the 'island' on a regular basis
5. Ocean (background): Represents the wider social and material environment in which the family is inserted
6. Clouds: Symbolise the challenges faced by the family
7. Boats: Represent the different institutions involved (school, health, social action), each with its own perspective and approach
8. Lighthouse: Symbolises B.'s team, offering guidance and support
9. Connecting lines: These show the interactions between the family and the different elements of the system

This visual map helps us to see:

1. The centrality of the family in the system
2. The multiple influences and stakeholders involved
3. The complexity of the relationships between the different elements
4. The crucial role of the mother as the 'interface' of the family
5. The unique position of the Neighbourhood Team as a source of guidance

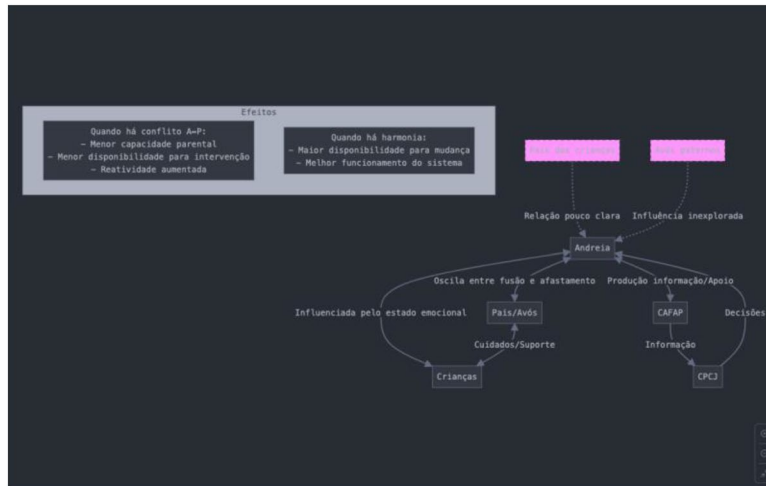
This visual representation captures the complexity of the system in a way that can reveal new perspectives or areas of potential intervention

Are there any elements you would like to add, remove or modify in this updated map? Does this visualisation help identify new areas for exploration or intervention?

Illustrative excerpts for Case 6: ext.user#4

Illustration of a Synthesis following Dialogical Scaffolding for the property of Relationality/Relations as Entities

AI/Claude Case Relationship Dynamics Map



[PHASE 1—STEP 1.1 Continuing]

I've created a visual representation of the relationships in Case, showing:

Core relationships between Andreia, her parents, and the children

- Institutional relationships with CAFAP and CPCJ

- Unexplored relationships (dotted lines)

- Key effects of relationship dynamics

Would you like me to continue with the previous question about how these relationships affect individual elements, or would you prefer to discuss what you observe in this visualization?

Illustrative example of a scaffolding modulating question on the property of Relationality/Relations as Entities

User We realise that the children are affected by the dynamics related to Andreia and her parents, we don't know if they are afraid, but we imagine that they are, because they are sheltered by the impact of these relationships. Children are most strongly affected by relational patterns. When the interaction is negative, Andreia no longer feels competent

[PHASE 1—STEP 1.1 Continuing]

Thank you for those important remarks. Let's explore the last fundamental question of this series:

[SC.RL.R.F.Q11] How can we continue to explore and update our understanding of relationships within the Case? Specifically:

What methods can we use to track and analyse changes in relationships over time?

How can we monitor the impact of these relationships on the children?

How can we assess the development of Andreia's parenting skills in relation to these dynamics?

AI/Claude Can you give us suggestions for that understanding of relations within the case?

AI/Claude [PHASE 1—STEP 1.2]

Before offering suggestions, let's explore a different dimension of relationships through a modulating question:

[SC.RL.R.MOD.SV.Q1] How do the relationships in Case change when we consider different contexts or environments? For example:

- What are relationships like at home vs. in the context of CAFAP?

- How do dynamics vary in times of stress vs. times of calm?

- What contextual factors seem to most influence the nature and strength of these relationships?

Once we've explored this, I'll be able to offer more specific suggestions for understanding

Note: This text was translated from Portuguese to English

Acknowledgements Filipe Santos was funded by an Individual Call to Scientific Employment Stimulus, awarded by the Portuguese Foundation for Science and Technology (CEEIND/03932/2017). Raquel Ribeiro was funded by the Scientific Employment Support Program (DL57/2016/CP1341/CT0030) of the Portuguese Foundation for Science and Technology (FCT). This study was funded by scientific sponsorship of Stefan Pernar, as part of the Complex Thinking & AI project. The requirements and challenges for the performance of complex thinking were previously presented in a poster presentation at the Satellite meeting of the Conference on Complex Systems "Complex Systems and Collective Intelligence", University Exeter, 2nd September, 2024.

Author contribution The authors adhere to an adapted version of the CRediT author statement (Allen, O'Connell, and Kiermer 2019) to recognise individual contributions: Ana Teixeira de Melo: Conceptualisation; methodology; software; data analysis; investigation; data curation; writing-original draft (major); visualisation; supervision; project administration; funding acquisition; writing-review of manuscript (major) Letícia Renault: Conceptualisation; methodology, software; data analysis; investigation; data-curation; writing-original draft (major); Leo Caves: Methodology; Software; data analysis; investigation; writing-original draft (major); visualisation; writing-review of manuscript (major) Philip Garnett: Software; investigation; writing-original draft (minor); writing-review of manuscript (minor) Paula Duarte Lopes: Investigation; Raquel Ribeiro: Investigation; writing-review & editing; Filipe Santos: Investigation; writing-review & editing.

Funding Open access funding provided by FCTIFCCN (b-on). Stefan Pernar, Scientific Sponsorship, Scientific Sponsorship, Scientific Sponsorship, Scientific Sponsorship, Scientific Sponsorship, Scientific Sponsorship, Foundation for Science and Technology, Scientific Employment Support Program: DL57/2016/CP1341/CT0030, Individual Call to Scientific Employment Stimulus: CEEIND/03932/2017.

Data availability Data referring to this study (e.g. logs of chats, transcriptions of sessions, databases of questions, transcripts of interviews, history of protocols) as well as the final suite of protocols can be found in the community Complex Thinking & AI at the Zenodo repository (<https://zenodo.org/communities/ctai>) and Github under the Complex Thinking Initiative (<https://github.com/complexthinking>) in the specific repository for the Complex Thinking & AI project (<https://github.com/complexthinking/complexthinkingai>).

Declarations

Conflict of Interests The authors declare no competing interests.

Open Access This article is licensed under a Creative Commons Attribution 4.0 International License, which permits use, sharing, adaptation, distribution and reproduction in any medium or format, as long as you give appropriate credit to the original author(s) and the source, provide a link to the Creative Commons licence, and indicate if changes were made. The images or other third party material in this article are included in the article's Creative Commons licence, unless indicated otherwise in a credit line to the material. If material is not included in the article's Creative Commons licence and your intended use is not permitted by statutory regulation or exceeds the permitted use, you will need to obtain permission directly from the copyright holder. To view a copy of this licence, visit <http://creativecommons.org/licenses/by/4.0/>.

References

- Ashby WR (1958) Requisite variety and its implications for the control of complex systems. *Cursos Congr Univ Santiago De Compostela* 1:83–99. https://doi.org/10.1007/978-1-4899-0718-9_28
- Baena-Rojas JJ, Ramírez-Montoya MS, Mazo-Cuervo DM, López-Caudana EO (2022) Traits of complex thinking: a bibliometric review of a disruptive construct in education. *J Intell* 10(3):37. <https://doi.org/10.3390/jintelligence10030037>
- Bai, Y., Kadavath, S., Kundu, S., Askill, A., Kernion, J., Jones, A., Chen, A., Goldie, A., Mirhoseini, A., McKinnon, C., Chen, C., Olsson, C., Olah, C., Hernandez, D., Drain, D., Ganguli, D., Li, D., Tran-Johnson, E., Perez, E., ... Kaplan, J. (2022). Constitutional AI: Harmlessness from AI Feedback. In *arXiv [cs.CL]*. arXiv. <http://arxiv.org/abs/2212.08073>
- Bartsch, A., Neuberger, C., Stark, B., Karnowski, V., Maurer, M., Pentzold, C., Quandt, T., Quiring, O., & Schemer, C. (2024). Epistemic authority in the digital public sphere: An integrative conceptual framework and research agenda. *Communication Theory*, qtae020. <https://doi.org/10.1093/ct/qtae020>
- Bateson G (1979) *Mind and nature: a necessary unity*, vol 255. Bantam Books, New York
- Beghetto RA (2023) A new horizon for possibility thinking: A conceptual case study of Human × AI collaboration. *PSS* 1(3):324–341. <https://doi.org/10.1177/27538699231160136>
- Brusilovsky P (1999) Adaptive and intelligent technologies for web-based education. *Ki* 13(4):19–25
- Caves, L., & Melo, A. T. (2018). (Gardening) Gardening: A Relational Framework for Complex Thinking About Complex Systems. In R. Walsh & S. Stepney (Eds.), *Narrating Complexity* (pp. 149–196). Springer International Publishing. https://doi.org/10.1007/978-3-319-64714-2_13
- Cowls J, Tsimados A, Taddeo M, Floridi L (2021) A definition, benchmark and database of AI for social good initiatives. *Nat Mach Intell* 3(2):111–115. <https://doi.org/10.1038/s42256-021-00296-0>
- Cui H, Yasseri T (2024) AI-Enhanced Collective Intelligence. *Patterns* 5(11):101074. <https://doi.org/10.1016/j.patter.2024.101074>
- Ferrario A, Facchini A, Termine A (2024) Experts or authorities? The strange case of the presumed epistemic superiority of artificial intelligence systems. *Minds Mach* 34:30
- Fiedler-Ferrara, N. (2010). *The Complex Thinking: Building of a New Paradigm*. V/RUS 3. http://www.nomads.usp.br/virus/virus03/PDF/review/1_en.pdf
- Fricker, M. (2007). *Epistemic injustice: Power and the ethics of knowing*. Oxford University Press.
- Google (2024). Prompting Guide 101: A Quick-Start Handbook for Effective Prompts.
- Hemmer P, Schemmer M, Vössing M, Kühl N (2021) Human-AI complementarity in hybrid intelligence systems: A structured literature review. *Pacific Asia Conf Informat Syst* 78:118
- Henrickson L, Meroño-Peñuela A (2023) Prompting meaning: a hermeneutic approach to optimising prompt engineering with ChatGPT. *AI & Soc*. <https://doi.org/10.1007/s00146-023-01752-8>
- Holton D, Clarke D (2006) Scaffolding and metacognition. *Int J Math Educ Sci Technol* 37(2):127–143. <https://doi.org/10.1080/00207390500285818>
- Jarrah MH, Lutz C, Newlands G (2022) Artificial intelligence, human intelligence and hybrid intelligence based on mutual augmentation. *Big Data Soc* 9(2):205395172211428. <https://doi.org/10.1177/20539517221142824>
- Jong, T. (2005). Scaffolding inquiry learning: How much intelligence is needed and by whom? *International Conference on Artificial Intelligence in Education*, 4. <https://www.semanticscholar.org/paper/Scaffolding-inquiry-learning%3A-How-much-intelligence-Jong/c18e7acaf9f7c2fc31c4dce9f716b7e811a5fce>

- Kamar, E. (2016). Directions in hybrid intelligence: Complementing AI systems with human intelligence. *International Joint Conference on Artificial Intelligence*, 4070–4073.
- Malone, E., Afroogh, S., DCruz, J., & Varshney, K. R. (2024). When trust is zero sum: Automation threat to epistemic agency. In *arXiv [cs.CY]*. arXiv. <http://arxiv.org/abs/2408.08846>
- Mancilla García M, Hertz T, Schlüter M, Preiser R, Woermann M (2020) Adopting process-relational perspectives to tackle the challenges of social-ecological systems research. *Ecol Soc* 25(1):29. <https://doi.org/10.5751/ES-11425-250129>
- Manyika J (2022) Getting AI right: Introductory notes on AI & society. *Daedalus* 151(2):5–27. https://doi.org/10.1162/daed_e_01897
- Marcus G, Davis E (2021) Insights for AI from the human mind. *Commun ACM* 64(1):38–41. <https://doi.org/10.1145/3392663>
- Melo AT (2020) Performing complexity: building foundations for the practice of complex thinking. Springer. <https://doi.org/10.1007/978-3-030-46245-1>
- Melo AT (2022) Conceptualizing and Designing (Co) Augmented Intelligence (s) Enacting Complex Thinking: Morphological Constraints Challenges and Implications. *Proceedings* 81(1):1–5. <https://doi.org/10.3390/proceedings2022081085>
- Melo AT, Renault R (2023) Complex thinking choreographies and transformative processes. Applying Education in a Complex World: Teaching and Learning. *AMPS Proceedings Series* 33.1.:247–258
- Melo AT, Renault L (2025) Complexigraphy: theoretical foundations and methodological challenges of mapping complex thinking. *Phenomenol Cogn Sci*. <https://doi.org/10.1007/s11097-025-10071-w>
- Melo AT, Caves LSD, Dewitt A, Clutton E, Macpherson R, Garnett P (2019) Thinking (in) complexity: (In) definitions and (mis)conceptions. *Syst Res Behav Sci* 37(1):154–169. <https://doi.org/10.1002/sres.2612>
- Mills, A. J., Durepos, G., & Wiebe, E. (2010). Exploratory case study. In A. J. Mills, G. Durepos, E. Wiebe (Eds.) *Encyclopedia of case study research* (Vol. 0, pp. 372–373). SAGE Publications, Inc., <https://doi.org/10.4135/9781412957397.n139>
- Modgil, S. (2018). Dialogical scaffolding for human and artificial agent reasoning. *5th International Workshop on Artificial Intelligence and Cognition, AIC 2017, 2090*, 58–71.
- Morales JAF (2020) Pensamiento Complejo: Una Revisión Sistemática de Artículos Científicos Indexados en Scopus 2016-2019. *Phainomenon* 19(2):303–323
- Morin E (2014) Complex Thinking for a Complex World—about Reductionism, Disjunction and Systemism. *Systema: Connect Matter Life Culture Technol* 2(1):14–22
- Morin E (2002) Seven complex lessons in education for the future. UNESCO
- Morin, E. (2005/1990). *Introduction à la pensée complexe*. Editions du Soleil.
- Naveed, H., Khan, A. U., Qiu, S., Saqib, M., Anwar, S., Usman, M., Akhtar, N., Barnes, N., & Mian, A. (2023). A comprehensive overview of large Language Models. In *arXiv [cs.CL]*. arXiv. <http://arxiv.org/abs/2307.06435>
- Newen, A., Bruin, L. D., & Gallagher, S. (Eds.). (2018). *The Oxford Handbook of 4E Cognition*. Oxford University Press.
- Nielsen J (1993) Iterative user-interface design. *Computer* 26(11):32–41. <https://doi.org/10.1109/2.241424>
- Olivo-Montañón PG, Sanabria-Z J, Molina-Espinosa JM, Quintero-Gómez L, Velarde-Camaquí D, Sánchez-Salgado LA, Gonzalez-Mendoza M, Breda A, Morales-Maure L, Alvarez-Icaza I (2024) MAICC model: development of complex thinking through citizen science project evaluation. *Front Edu* 9:1392104. <https://doi.org/10.3389/educ.2024.1392104>
- OpenAI. (2023). GPT-4 Technical Report. <https://cdn.openai.com/papers/gpt-4.pdf>.
- Patiño A, Ramírez-Montoya MS, Ibarra-Vazquez G (2023) Trends and research outcomes of technology-based interventions for complex thinking development in higher education: A review of scientific publications. *Contemp Edu Technol* 15(4):ep447
- Patton, M. Q. (2011). *Developmental Evaluation: Applying Complexity Concepts to Enhance Innovation and Use*. Guilford Press.
- Ramírez-Montoya MS, Álvarez-Icaza I, Sanabria-Z J, Lopez-Caudana E, Alonso-Galicia PE, Miranda J (2021) Scaling complex thinking for everyone: a conceptual and methodological framework. In: Ninth international conference on technological ecosystems for enhancing multiculturalism (TEEM'21), pp 806–811. <https://doi.org/10.1145/3486011.3486562>
- Ramírez-Montoya MS, Castillo-Martínez IM, Sanabria-Z J, Miranda J (2022) Complex thinking in the framework of education 4.0 and open innovation—a systematic literature review. *J Open Innovat: Technol Market Complex* 8(1):4. <https://doi.org/10.3390/joimc8010004>
- Ray PP (2023) ChatGPT: A comprehensive review on background, applications, key challenges, bias, ethics, limitations and future scope. *Internet Things Cyber-Phys Syst* 3:121–154. <https://doi.org/10.1016/j.iotcps.2023.04.003>
- Renn O (2017) Risk governance: coping with uncertainty in a complex world. Routledge
- Rogers K, Luton R, Biggs H, Biggs R, Blignaut S, Choles AG, Palmer C, Tangwe P (2013) Fostering complexity thinking in action research for change in social-ecological systems. *Ecol Soc* 18(2):31. <https://doi.org/10.5751/ES-05330-180231>
- Roose, K. (2024, December 13). Why Anthropic's Claude Is a Hit with Tech Insiders. *The New York Times*. <https://www.nytimes.com/2024/12/13/technology/claude-ai-anthropic.html>
- Sætra HS (2022) Scaffolding human champions: AI as a more competent other. *Human Arenas*. <https://doi.org/10.1007/s42087-022-00304-8>
- Saldaña J (2016) The coding manual for qualitative researchers. Sage
- Salomon G, Perkins DN, Globerson T (1991) Partners in cognition: Extending human intelligence with intelligent technologies. *Educ Res* 20(3):2–9. <https://doi.org/10.3102/0013189x020003002>
- Saravia, E. (2022). Prompt Engineering Guide. <https://github.com/dair-ai/prompt-engineering-guide>. <https://www.promptingguide.ai/>
- Sarrion, E. (2023). Basic Rules for Asking Questions to ChatGPT. In E. Sarrion (Ed.), *Exploring the Power of ChatGPT: Applications, Techniques, and Implications* (pp. 93–98). Apress. https://doi.org/10.1007/978-1-4842-9529-8_10
- SEBoK. (2024). System-of-Interest (glossary). SEBoK. Guide to Systems Engineering Body of Knowledge. [https://sebokwiki.org/wiki/System-of-Interest_\(glossary\)](https://sebokwiki.org/wiki/System-of-Interest_(glossary))
- Silva Pacheco C, Iturra Herrera C (2021) A conceptual proposal and operational definitions of the cognitive processes of complex thinking. *Think Skill Creat* 39:100794. <https://doi.org/10.1016/j.tsc.2021.100794>
- Stetsenko A (2017) The Transformative Mind: Expanding Vygotsky's Approach to Development and Education. Cambridge University Press
- The Open University. (2023). *Systems engineering: Challenging complexity*. The Open University.
- Tobón S, Luna-Nemecio J (2021) Complex thinking and sustainable social development: validity and reliability of the COMPLEX-21 Scale. *Sustainability: Sci Pract Policy* 13(12):6591
- Varela, F., Thompson, E., & Rosch, E. (2016). *The Embodied Mind. Cognitive Science and Human Experience* (Rev. Ed). Massachusetts Institute of Technology.
- Varga, P. L. (2018). Mixed methods research: a method for complex systems. In E. Mitleton-Kelly, A. Paraskevas, & D. C. (Eds.), *Handbook of Research Methods in Complexity Science* (pp. 545–566). Edward Elgar Publishing. <https://doi.org/10.4337/9781785364426.00038>

- Velásquez-Henao JD, Franco-Cardona CJ, Cadavid-Higueta L (2023) Prompt Engineering: a methodology for optimizing interactions with AI-Language Models in the field of engineering. *Dyna* 90(230):9–17
- Vygotsky, L. S. (1978). *Mind in Society: Development of Higher Psychological Processes*. Harvard University Press.
- Wells J (2013) Complexity and sustainability. Routledge
- White, J., Fu, Q., Hays, S., Sandborn, M., Olea, C., Gilbert, H., Elnashar, A., Spencer-Smith, J., & Schmidt, D. C. (2023). A prompt pattern catalog to enhance prompt engineering with ChatGPT. In *arXiv [cs.SE]*. arXiv. <https://doi.org/10.48550/arXiv.2302.11382>
- Wood D, Bruner JS, Ross G (1976) The role of tutoring in problem solving. *J Child Psychol Psychiat* 17(2):89–100. <https://doi.org/10.1111/j.1469-7610.1976.tb00381.x>
- Ye, Q., Axmed, M., Pryzant, R., & Khani, F. (2023). Prompt Engineering a Prompt Engineer. In *arXiv [cs.CL]*. arXiv. <http://arxiv.org/abs/2311.05661>
- Yin, R. K. (Ed.). (2018). *Case study research and applications: design and methods* (6th Ed.). Sage.
- Zamfirescu-Pereira, J. D., Wong, R. Y., Hartmann, B., & Yang, Q. (2023). Why Johnny can't prompt: How non-AI experts try (and fail) to design LLM prompts. *Proceedings of the 2023 CHI Conference on Human Factors in Computing Systems*, 1–21. <https://doi.org/10.1145/3544548.3581388>

Publisher's Note Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.