31 July 2022

WHITE PAPER

*A*I?

Human-AI Collaboration while Keeping a Sense of Self

Self-Awareness and AI in Mixed Reality
A Serious Game

by
Ida Helena Rust

Dear highly knowledgeable reader with a background in computer science,

I am not asking you not to be critical or even skeptical. I even invite you to confront me with your critical remarks, because only then can my work possibly have any value for others.
Yet, I would like to ask you to choose a personal memory that you believe defines who you are today. Recall that moment, feel that moment and ask yourself, “Who am I without this experience?”

The author

CONTENTS

Summary …………………………………………………………………………………… 4

1 Introduction ………………………………………………………………………………… 5
- The current approach to Human-AI collaboration is not a human solution
- Problem description
- Proposed solution
- ***AI?***'s unique features
- Mixed reality may blur human reflection
- A broader view on reality
- Content description

2 **Human-in-the-Loop AI automation**……………………………………………….......... 8
- Problem description
- What is AI?
- How AI forms a system
- Closing the loop with the human
- Optimizing the professional for system integration
- The numerical human
- Improved trustworthiness of humans by means of affective computing
- Human-AI collaboration and trust
- Computational thinking

3 **The Human behind the Professional**……………………………………………….. 11
- The self as reason
- The self as scientific object

3.1 **Phenomenology**…………………………………………………………………………… 11
- The world does not exist apart from the experience
- Prioritizing the individual's experience
- Science and Phenomenology coexist
- Self-awareness
- Pre-reflective self-awareness

3.2 **Post-phenomenology** ……………………………………………………………… 12
- Technology creates new realities
- A Post-phenomenological take on Human-AI trust
- The human does not exist independent from AI
- Human-Technology relations
- Human-AI relations
- I amidst of AI

4 ***AI?* a Serious Game**……………………………………………………………………. 15
- What is ***AI?***

4.1 **Serious Gaming**……………………………………………………………………….... 16
- Game-based learning
- Game mechanics
- Visual aesthetic, sound and narrative design
- Avatars and digital twins
- Incentive system
- Intrinsic motivations
- Personalization of the game
- Content and Skills

4.2 **Serious Gaming in Mixed Reality**…… 20
- Mixed reality
- Mixed reality in professional environments
- Serious Gaming in Mixed Reality
- Ubiquitous learning and computational thinking

5 ***AI?* Applied: The Medical Specialist**…… 22
- AI in medicine
- AI in the operating room of the future
- Mixed reality in the OR
- Trust in AI in medicine
- Reflecting on the trust-relation with AI
- ***AI?*** applied
- The integration of ***AI?*** in the work environment

6 **Concluding Thoughts**…… 26
- A broader application of ***AI?***
- Development challenges and further research

Acknowledgements …… 27
About the author …… 27
References …… 28
Appendix: Supplementary information…… 32

SUMMARY

This white paper is directed at AI companies that focus on seamless Human-AI collaboration, and organizations that use AI and strive to maintain meaningful work and remain a human organization rather than a technology-driven automated environment.
The paper presents a rudimentary non-technical outline for a mixed reality AI-application, hereafter referred to as ***AI?***, that uses serious gaming to train domain experts in complex and highly automated environments, to collaborate with artificial intelligence (AI) from a sense of self.[1] The self, as presented, consists of three elements: (1) the Thinking Self, (2) the Experiential Self, and (3) the Composite Self, i.e. an understanding of self that sees AI as a part of oneself. Self-awareness enables the domain expert to reflect on different alternatives from a human/system perspective and to make human-beneficial decisions based on the possibilities offered by AI.

The paper takes the collaboration between humans and AI as its starting point. From there it explains that in order to collaborate fruitfully with AI, the domain expert should not only use their professional knowledge and predefined skillset, but also their individual experience. ***AI?*** enables organizations to be human-driven, and not driven only by computational efficiency and effectiveness as the sole means to organizational success.
The serious game ***AI?*** consists of four levels:

Level 1	Computational Thinking: i.e. learning to understand how AI perceives and engages with the world.
Level 2	Mindfulness: becoming experientially self-conscious[2].
Level 3	The Composite Self: understanding the relationship between oneself and AI, and the possibilities that arise from it
Bonus Level	This level carries the name 'Kill the Buddha', referring to the Buddhist expression[3], which in this context means to get rid of ***AI?*** if the tool becomes the guiding principle for the user's intentionality.

The paper combines a philosophical reflection on the future role of domain experts in professional AI-automated environments with current research and development in AI. Phenomenology and Post-phenomenology will be discussed as an alternative for the scientifically objectified view of the human and the world. The problem the paper identified, is that humans may become rational agents in a computational-rational system, which leads to the elimination of human-driven intentionality and agency. ***AI?*** intends to offer a solution by creating space in the cyber-physical-human system for human intentionality as the result of self-awareness.
AI? seeks to go beyond understanding the human being as a set of measurable properties. ***AI?*** does not provide insight into the psychological mechanisms behind human experiences and emotions, and dismisses the collection and sharing of data with third parties about internal human states, progress in the game process and created scenarios, unless the user gives permission.
AI? differs from current Human-AI collaboration approaches, in (a) that it makes the computational rationality of AI explicit, instead of anthropomorphizing its thinking; (b) by combining the practice of mindfulness with a critical reflection on the Human-AI relationship; and (c) by putting the intentionality of humans first, rather than technologically accommodating humans to act according to AI's computational rational intentions.

[1] This white paper anticipates that mixed reality will be the future and game-based learning will be widely adopted as an educational tool.
[2] The terms 'self-awareness' and 'self-consciousness' will be used interchangeably, in example of [30].
[3] A Koan ascribed to the Chinese Ch'an master Linji Yixuan (d. 866), which start with: "If you meet the Buddha, kill the Buddha". The general interpretation of this Koan is, that we have Buddha nature within ourselves and that if something outside claims to be the Buddha, it would mean that we have turned our backs on our own nature to follow an idol.

1. INTRODUCTION

The future is Human-AI collaboration. AI-based automation and devices are becoming mainstream in organizations. Organizations in the process of digitalization are facing questions, such as: how will the work culture change, which functions will disappear and be created, should AI be part of shared decision making, and what is the influence of AI on human agency and control [61].

The current approach to Human-AI collaboration is not a human solution

Professionals may feel anxious, suspicious, overwhelmed, inadequate, among other sentiments, about the ongoing AI automation in the workplace [66]. These feelings lead to a lack of trust and can hinder successful collaboration with AI. The felt disconnection with AI is currently being resolved through the development of transparent AI and affective computing. However, this solution approaches the problem from a technological point of view, rather than providing a space for intrinsically driven human participation.

Problem description

Concerns have been raised that AI is being developed in such a way, that its bias goes unnoticed in decision-making and that its decisions will be unintelligible due to the amount of data and the complexity of the analyses involved, which will lead to mistrust.[4] It is also noted that too much reliance on AI can lead to the loss of key skills, expert supervision, professional identity and psychological well-being [52], [59].

This white paper argues that in the future, domain experts with their unique experiences and insights may become part of the AI automation system and thus can be reduced to optimizable rational agents. This development would prevent human-driven progressive changes to occur. Moreover, affective computing is able to intercept emotions and can decide to interfere with them, without humans being aware of it. Thus being able to manipulate human intentionality and agency.

It is becoming clear that AI's understanding of the world based on models and categorizations is far from flawless. Moreover, by pre-defining and controlling desired outcomes, on human insight and experience based alterations are rejected or not possible at all. Therefore, it is recommendable that a middle ground is sought, where the with AI-interacting human recognizes the benefits of AI, but is given a superior role in a human-driven AI automated[5] organization. By ‘Human-driven’ is meant here:

1) Intrinsic human worth
2) Decision-making and the responsibility of these decisions are in human hands
3) Decisions serve the human

Proposed solution

The proposed solution is ***AI?***, an AI application, integrated in mixed reality, that teaches in a game-inspired format: Computational Thinking, Mindfulness and Post-phenomenological Self-Awareness, called ‘The Composite Self’: i.e. understanding that one should not see oneself independent from AI and that intentionality and agency emerge because of that.

The paper posits that in order to become an authentic human-centered AI-automated organization, the human must first become aware of the respective AI's understanding of the world. Secondly, become experientially aware of oneself. Here, the learner is guided in an exercise of experiential self-awareness. Thirdly, understand their relationship with the respective AI and realize that AI cannot be seen independently of them. The professional learns to bring together their experiential self-awareness and reasoning skills, to reflect on how the respective AI influences one’s intentionality. Subsequently, the learner will explore different alternatives and ultimately choose the most optimal alternative from a human/system point of view. A fourth bonus level, allows the user to critically assess ***AI?*** and offers the option to dismiss it.

The goal is to establish an internal human anchor point from which decisions and actions emerge that can be executed by AI, humans or in collaboration.

---

[4] Other concerns include accountability, governance and privacy issues [24].

[5] Instead of the common terms ‘smart automation’ and ‘intelligent automation’, the term AI automation is used to indicate artificial intelligent automation.

AI? unique features

AI? cannot be substituted by a standalone computational thinking or mindfulness training application, because ***AI?*** brings both computational thinking and mindfulness together as human aspects and uses this awareness to consciously explore the possibilities that AI offers. The justification for understanding computational thinking first and then practicing mindfulness is because the goal of ***AI?*** is to improve the collaboration between humans and AI. Understanding how a computer thinks, directly benefits the organization by making AI less perplexing for the professional, therewith increasing trust. The importance of computational thinking is evident from its current implementation as part of the educational curriculum. In addition, there is also a strong case for transparent AI to improve the user's relationship with AI.

The term 'Mindfulness' is associated with the increasingly popular practice of reducing stress.[6] In the case of ***AI?***, the term denotes the act of becoming experientially aware of one's presence, complementary to our ability to think consciously. Unlike mindfulness applications that have a therapeutic or spiritual purpose, ***AI?*** takes into account the automated environments in which the domain expert operates. By combining computational thinking with mindfulness, ***AI?*** supports the professional in returning to their experiential self in the midst of the automation. From that sense of self ***AI?*** guides the domain expert in a conscious relationship with AI technology.

AI? is a 'serious game'. Serious gaming combines educational methodology with game elements. Over the years, serious gaming has become an increasingly popular and effective means to train professionals. Especially in environments that are subject to high stake unpredictable and changing conditions. Besides, because ***AI?*** cannot be imposed and can only be used based on intrinsic motivation, the game form makes voluntary engagement more attractive.

Mixed reality may blur human reflection

AI? is envisioned as an AI application that runs in a head-mounted display and identifies other AI in mixed reality.[7] The paper anticipates the increasing use of augmented reality (AR) and virtual reality (VR) in general, and professional environments in particular.[8] Mixed reality is an advanced form of AR in which physical reality and VR overlap and interact. It can be expected that in the not too far future, domain experts will permanently use head-mounted displays and work in a mixed reality, while interacting with virtual objects and information.

When mixed reality becomes the 'new reality', it will be more difficult to distance oneself from the technologized environment and be aware of one's own thoughts and feelings. This situation may lead to uncritical engagement with AI and the loss of human input.

A broader view on reality

The content of the paper is partly philosophical in nature. AI perceives the world computationally rationally. This is an abstraction of reality. In order to have a more complete perception of reality and to be able to shape a human desirable future, it is important to acknowledge the human experience of reality, in addition to the scientific objective understanding. Phenomenology takes human experience as the starting point for our engagement with the world. Post-phenomenology, in turn, examines how technology changes the human experience of reality and approaches AI as an inextricably integrated part of human life.

The paper takes a Phenomenological stance on human consciousness to bridge the divide between reason and feeling and close the gap between humans and AI. In doing so, it aims to show that there is a way of 'looking at ourselves' outside the scientific paradigm. Phenomenology studies the subjective experience of an individual without separating the

[6] According to Future Market Insights, the global mindfulness meditation application market is expected to double by 2030.

[7] Since mixed reality has not taken yet off at large, In the short term, ***AI?*** can also be developed for handheld devices.

[8] According to a white paper on immersive media technologies of the World Economic Forum, the global market growth estimations are based on Virtual and Augmented Reality. The VR market is expected to grow to $20.9 billion by 2025, and AR to $77.0 billion by 2025.

human from what they experience. Self-consciousness, understood Phenomenologically, is an experiential state prior to thinking about what one actually experiences.
Current research and development (R&D) in the field of human-AI collaboration approaches humans from a scientific point of view. This assumes that humans remain unchanged in their interaction with AI. On the other hand, by seeing the human as an object, the human becomes subject to possible alterations by humans themselves or by AI. These alterations should be a conscious individual choice and not a decision imposed to optimize system performance. That said, Post-phenomenology argues that technology changes human intentionality and that this should be assumed for understanding AI and exploring our future and making human-beneficial decisions with AI.

Level 3 of ***A*I?**, the Composite Self, invites the learner into a Post-phenomenological reflection on the Human-AI relation. The three key aspects of this reflection are:

1) Becoming aware that the human does not exist independent from AI and that both mutually shape each other.
2) Becoming aware that AI has a model of the world, based on what can be measured (ideally a scientific understanding) and does not necessarily resembles the reality as as an individual sees it.
3) Becoming aware that the user has a perception of the reality, which is partly formed by their experiences.

The insights resulting from this reflection enables creative advancement in which humanity resides.

Content description
The paper will first sketch an image of what AI is, how it forms a system and how computational rational thinking optimizes the professional for system integration. Then, it will briefly describe Phenomenological and Post-phenomenological theory to provide an additional non-scientific view on the human and the world, and their relation with AI technology. Part four provides a first outline of what ***A*I?** as a serious game may look like based on serious game theory and examples; and will say something about mixed reality and serious gaming in mixed reality. Finally, it will apply the initial game design to the environment of the medical specialist of the future, to explore how ***A*I?** will function in this particular work environment. The paper ends with some concluding thoughts on the wider application of ***A*I?** and development challenges.

This white paper does not provide an ultimate development guide for ***A*I?**. What it offers, ahead of impactful technology trends, is a defense and non-technical rudimentary guidelines for developing an AI application that puts humans at the center of an organization. The paper also has the ambition to raise awareness of the implications of the computational rationalization of reality and to stimulate more research into preserving the immeasurable human rather than abstracting them into their digital twin.

2 HUMAN-IN-THE-LOOP AI AUTOMATION

Problem description

The numerical, abstracted human as understood by AI, will become another rational agent, i.e. a complex function that can be optimized by AI to achieve a goal. The question is: who or what decides the goal? If humans only depend on their reason, then the computational rational system will be reinforced. If the computational-rational model of reality, monitored and controlled by AI, becomes the reference, then the unique characteristics of the individual human that cannot be grasped by AI, or are considered unwanted, or not needed, or both, will be eliminated. As a consequence, the sense of an experiential self will be lost.

One aspect of being human, in addition to having reason, is the unique emotional experience of reality. It is undeniable that emotions, playing an important role both in decision making, and goal setting, can be intercepted and manipulated by AI. Therefore, human emotions that may lead to choices and objectives not in alignment with the goals AI is trained to accomplish, will be intercepted and manipulated by AI (as well as the environment that may evoke a certain emotion). This will lead to a further decoupling of the 'feeling human' and their reason in furtherance of reason (reinforcing the computational rational model of the world).

What is AI?

The term AI harbors the ambition to create an artificial equivalent of human intelligence. In this white paper, the term refers to computational-rational agents. AI consists of a set of algorithms. An algorithm is a sequence of instructions that always produces the same result, given the same input [20]. AI is designed to achieve the best possible desired performance, based on identifying and approximating the best actions and trade-offs, which are subject to optimization, in a given environment that provides the digital input for the algorithms [10],[32], [20],[69]. Today's advanced forms of AI are machine learning and deep learning. Machine learning automates complex cognitive tasks that humans would otherwise do, such as pattern recognition, statistical modelling, data mining, knowledge discovery, and predictive analytics [20]. Deep learning, as a form of machine learning, creates stacking layers of data transformation, to refine and optimize its own prediction or categorization, representing them as referential data sets to optimize the output function with minimal or no human intervention. Once the machine learning algorithm has achieved the desired result, as is currently determined by humans, it is applied in practice as a very complex mathematical calculation.

How AI forms a system

AI is developed to achieve a certain goal and to optimize the path to achieve this goal. AI seeks control over its environment in order to function optimally. It is essential for AI to have the availability over as much data as possible to know what is happening, and to be able to predict what will happen. AI does this on the basis of pattern recognition, i.e. looking for similarities with historical data and then deciding if something is 'good' because it meets predefined parameters. With this information, AI creates a model of the environment. This computational-rational understanding of the environment enables AI to accomplish its objective.

Connecting AI with other AI increases the amount of data that AI can use and make inferences from, and therewith possibilities for AI application. AI needs sensors to interact with the physical environment. These sensors, and the processed data from these sensors, can be used by other AI. The connected AI-powered objects and systems as such, form new and often larger AI systems.

Closing the loop with the human

The decision to whether or not to replace the human for AI is based on a rational estimation of the time and cost of the labor inputs and the quantity and quality of the outputs [41]. Processes in various professional domains are automated. The processes that are automated are physical repetitive tasks and the collection and processing of data. When AI is added to automation, more complex human cognitive interventions can also be performed by AI.

The term human-in-the-loop[9] is used for the purposeful assignment of specific functions and decision-making to humans in an automated system. An emerging area in AI automation is the development of cyber-physical-human systems. A cyber-physical-human system takes the human-machine interaction as its starting point. Functionalities are flexibly assigned to AI

[9] The opposite is closed-loop automation, which means that there is no human intervention required at all.

and humans, based on the probability of success of performing a specific task. In this hybrid setting, humans and AI support and adapt to each other.

Optimizing the professional for system integration
The increasing AI automation puts into effect, that professionals are expected to adapt by developing relevant skills and competencies[10] that cannot be automated. This specialization allows professionals to have more time for their predefined core tasks [41]. In this transition to cyber-physical-human systems, professionals are being described in computer-related terms [28]. Human performance is expressed in terms of value and efficiency. Terms like 'memory' and 'rationality' applied to humans, are nowadays computationally understood [2]. Computer terms as 'input' and 'processing' are applied to humans. And the term 'agile'[11] also has its origins in automation, which is aimed at professional improvement according to computational-rational standards.

The numerical human
Although some AI put humans at the center, they are in fact integrating humans into their systems thinking. The system accommodates the professional, aimed at an optimized rational result. AI creates and works with a simplified version of human reality. Since AI engages in pattern recognition and the abstraction of data, it will look for computable commonalities among humans.
AI understands humans as data. In addition to the intentional output that humans produce in the form of text and speech, AI can draw inferences about humans by recognizing face, voice, gait, heart print, accent, emotions, personality, bone, genetics, gesture, touch, and bioacoustics[12] based on measurability.
The by AI understood human is an abstracted version. The transition from the physical human to a numerical one results in information being lost or changed, which may have consequences for the human being in physical reality [8].

Improved trustworthiness of humans by means of affective computing
Predictability, shared context and mutual trust are crucial for successful Human-AI collaboration [67]. In order for AI to interact successfully (from a computational-rational perspective) with humans, it seeks insight into the processes underlying human actions. Fluctuations in emotional and mental states can affect a person's trustworthiness [34],[79]. It is therefore important for AI to recognize and understand what affects them, in order to intervene timely to make humans less unpredictable [5]. The means to do that is affective computing. Affective computing is the development of AI that is able to anticipate mental, emotional and physical human states by recognizing, analyzing and interpreting human physiology and language usage.

Human-AI collaboration and trust
Due to the increasing complexity of the analyses of AI, humans have to rely on trust in their interaction with AI. Trust means that there is no certainty about the outcome and that one consciously involves oneself in an individual risk.
AI uses the two following strategies to be perceived as trustworthy: (1) transparency: i.e. explaining how it arrives to a conclusion and providing insight into the data and processes with which it works [34],[5],[89],[25],[63], and (2) anthropomorphization. These strategies are based on two forms of trust: rational trust and emotional trust.[13]
Rational trust means being able to choose the best option for oneself, given a specific scenario and goal to achieve. The trustor estimates the likelihood of the trustee to perform a certain action that is beneficial for the accomplishment of a goal of the former [31],[79],[80],[27]. Providing insight and explanations can be used by humans to rationally estimate the trustworthiness of the AI. It is said that one cannot be convinced to trust. Human

---

[10] Critical thinking, creativity, communication and collaboration are, among others, identified as desired competencies in a semi-automated working environment.
[11] The word 'agile' derives its contemporary popularity from the software development manifest (2001) of the same name. The manifest presents a list of computational-rational requirements for humans that are necessary for successful software development, such as: goal-oriented self-development, self-organizing teams and permanent quality control [45].
[12] The Future Today Institute 's15th Anniversary Tech Trends Report 2022.
[13] Overall, three forms of trust can be distinguished: affective trust, normative trust and rational trust [44],[71]. The account of normative trust is grounded in rational trust. Kant uses this notion to explain moral faith in an impersonal God [26].

trust is often emotional and unconscious, based on goodwill and can be seen as an ‘attitude from the heart’ [6],[44],[15],[21],[26]. Therefore, the partial anthropomorphisation[14] of AI, i.e. adding human characteristics to a non-human entity, may evoke human affection for AI and therewith emotional trust [14].

However, the use of AI may lead to over-trust in AI [59]. In this case the professional uncritically adopts the information or recommendation of AI. AI may have an understanding of reality that does not correspond with what domain experts experience as being the case. This wrong image can be the result of bias creep, which means that the AI has created a model based on non-representative data, which leads to erroneous output. The non-representative data is the result of the cognitive biases of its developer or of training data sets or because it is incomplete.

Computational thinking

Computational thinking is at the heart of AI. The term 'computational thinking' is popularized by [90] in her eponymous publication. She argues that computational thinking is a skill that everyone should learn, because computational thinking will become an integrated part of human life. It should therefore be part of other academic disciplines, applied as algorithmic medicine, computational economics, digital humanities and many others [91].

Computational thinking implies efficient problem solving by means of the decomposition of complex problems, practical reasoning, thinking recursively, pattern recognition and abstracting data input, and aims at the optimization of task performance, among others. Computational thinking is how people think and not how computers think. Through computational thinking, humans are able to develop AI to create progress for humanity [90],[91],[74].

However, leading up to the next part, by making the future depend on computational thinking, it is the AI development itself that urges computational thinking.[15] Additionally, by putting human reason first, an important part of what constitutes the human is side-lined: i.e. the experiential human.

[14] … for too much human likeness may cause suspicion [34]. Therefore, a certain amount of revealing its ‘computerness’ is perceived as more objective [1],[34].

[15] As what [90] describes as an incestuous relation.

3 THE HUMAN BEHIND THE PROFESSIONAL

The self as reason
AI emanates from human thinking. Not only as technology, born of human intellectual ingenuity, but also in how it reasons. AI reflects humans' unique ability to reason and think logically. AI is based on the assumption that the human brain is an information processing device and that human reasoning can be understood as computation. As well as that reality can be comprehended empirically through deductive reasoning, where rules can be inferred from human experience. Descartes' famous phrase, "I think, therefore I am," heralded the equating of the self with their thinking and with that, the importance of human reason, which paved the way for the development of computers [22],[11],[69],[4].

The self as a scientific object
In general, in the Western tradition, the self is understood as our ability to reason and feel. Throughout history, the understanding of the self underwent several changes. In the early Christian tradition, the self was understood as the connection with God, driven by love and the desire for knowledge. From the Age of Enlightenment, the human began to see their conscious experiences as objects of thought, which can only be approached by reason to understand themselves.[16] At the end of the eighteenth century, the philosopher Immanuel Kant pointed out that the objective world is a prerequisite for conscious thought. A century later, with the advent of psychology, also emotions and behavior became subject of scientific inquiry. In recent decades, the understanding of consciousness has expanded further by understanding the mind as chemical and physical neural processes.
Reason objectifies the human mind in two ways: first, by making the mind into a comprehensible study object by different scientific means. Secondly, by making the mind the object of self-consciousness by the individual themselves [94],[97].

3.1 Phenomenology

The world does not exist apart from the experience
Phenomenology, with roots in both philosophy and psychology, points out that our conscious experiences of the world (phenomena) exist primarily because there is an experiencing human being [95],[30],[97],[73]. Phenomenology takes the inseparable and interdependent relation between the human mind, body and world as its starting point and inquires the structures of the individual's conscious experiences with the world.

Prioritizing the individual's experience
Phenomenology does not objectify or isolate the mind in a Cartesian way. Phenomenology is concerned with how things appear to an individual and how one perceives something. An individual's experience is the entry to what we can know [55] and Phenomenology does not look for similarities in the different individual experiences. Phenomenology is concerned with one's subjective relationship with the world and not with consciousness as the thinking mind. Thinking about oneself should be seen differently from experiencing oneself.

Science and Phenomenology coexist
Phenomenology does not oppose the worldview defined in scientific terms as objective properties, but defends that human reality cannot be fully understood by a purely scientific approach. Reality in its totality should not be understood in such a way that it fits into a quantifiable scientific understanding of the world [97]. Phenomenology revolves around the individual's experience of things and does not intend to proof that what is experienced, really exists.
The late biologist and neuroscientist Francisco Varela [82] argues that science would do well to include human experiences untouched by thought. Science is descriptive and pretends not to be involved, when in fact it is a human affair: "lived experience is where we start from and where all must link back to, like a guiding thread."[17] Moreover, humans naturally do not see

---

[16] Descartes (1640) as cited in [30]: "By the word 'thought', I understand all that of which we are conscious as operating in us".
[17] Varela 2003 in [16]: 120

the world as scientific parts, but as interconnected with a coherent meaning. From an experiential awareness, the self is a "stream of experience [of the world]. It is a process and not a thing" [82].

Self-awareness
Phenomenological self-awareness holds that one is aware of their experience, and that one can consciously choose to approach this experience as an object of thought [30]. In Phenomenology, being aware equates experiencing the world. Experiencing something is a self-manifestation [94]; and since we always experience something, that 'something' is part of our 'selves'.

Pre-reflective self-awareness
Pre-reflective self-awareness is the stage before any description or classification is given to that which is experienced. The human experience is not yet apprehended by the intellect [68]. Experienced emotions are in this stage not labelled by the person who experiences them, because that would involve thought.
Pre-reflective experiences are found in animals and children under the age of three. They have no awareness of their own mental states. They just experience [96],[30]. The adding of self-awareness to pre-reflective means that a person is aware of their experience, but hasn't made it into an object of thought. This is popularly referred to as 'Mindfulness' [16].[18]

3.2 Post-phenomenology

Technology creates new realities
Post-phenomenology builds on Phenomenology and inquires how technology[19], both intentionally and unintentionally, creates new ways for the individual to experience the world. Technology is responsible for a permanent change of the individual's perception of reality. For example, eyeglasses alter one's vision of the world and ultrasound changes the relationship of the parents with the unborn baby [85,[87].
According to Post-phenomenology, humans and technology co-shape each other [85],[87], [43],[3]. Humans create technology, which alters the way humans interact with the world, which in its turn leads to novel technology. As Homo faber (the working man) we create technology to increase our wellbeing, which at the same time stimulates humanity to create new inventions [43]. AI development as the product of human intelligence is essentially part of human survival strategy [10].

A Post-phenomenological take on Human-AI trust
Trust in AI from a Post-phenomenological perspective implies the acceptance of technology as an inevitable part of our existence. This acceptance, which is described as an act of surrender with confidence, results in a person taking responsibility for their relationship with technology. The trust is deepened by using the technology. Trust grows in the interaction, just like in a relationship between two people. Trust in AI-technology should therefore be seen as an evolving and continuous process [47].

The human does not exist independent from AI
To strengthen the trust-relation with AI, and enforce the trust that this relationship will develop in a direction desired by humans, it is important to transcend the subject-object relationship we presume to have with AI and to critically assess the relation 'from within' [87]. Here, the term 'critical' does not necessarily refer to 'critical thinking', as critical thinking equates to scientific and logical thinking [17],[18]. 'Critical' is to be understood as the awareness of that which does not meet individual expectations, and where subsequently the source of this mismatch of expectation is explored, both experientially and through reason. The insights that emerge from this exploration serve human decision-making.

[18] The author is aware of the research paper by [77], in which they state that mindfulness cannot be compared to phenomenological research. The author emphasizes that she does not equate mindfulness with Phenomenology, but does interpret, together with [16],[82], the state of 'pre-reflective self-awareness' as mindfulness.
[19] With technology is meant human-made tools in its broadest sense. The current association with technology as the application of scientific insight to manipulate nature to meet societal demand, rose during the industrial revolution.

Human-Technology relations

Human intentionality is generally understood as a conscious involvement with something in the world. To make this conscious decision, one must have a prior understanding of that what one wants to engage with. Technology mediates the human experience of the world and influences therewith human intentionality.

[42] Has identified four different intentional relations of humans with the world involving non-AI technology, as schematized in table 1 [42],[86].

Human – Technology Relations	
1. Embodiment relations	
Scheme	(Human – Technology) →[20] Environment
Description	The human together with a specific technology create a different engagement with the environment.
Example	Eyeglasses
2. Hermeneutic relations	
Scheme	Human → (Technology – Environment)
Description	The human interact with the environment as interpreted by technology.
Example	Thermometer
3. Alterity relations	
Scheme	Human → Technology (–Environment)
Description	The human interacts intentionally and directly with the technology, while the environment is in the background.
Example	ATM
4. Background relations	
Scheme	Human → (Technology –) Environment
Description	The human interacts with the environment, while technology is invisibly supportive in the environment.
Example	Air-conditioning system

Table 1

Human-AI relations

With the rise of AI, [86].[87] identified four more relations in which technology itself has become intentional (table 2). An accepted understanding of AI and intentionality is: "perform[ing] an action when it believes it will and also believes that it could do otherwise" [54]. The having of intentionality by AI, makes the human relationship with it different from other technologies. The consequences of this new kind of relationship with technology are the subject of current Post-phenomenological research.

A ninth relation with AI, named 'relegation', which has been identified by [88], anticipates 'intention(ally) influencing AI recommendation systems' in mixed reality. This relation is based on the artistic video work titled "Hyper-Reality" by Keiichi Matsuda.

While augmented reality "adds [indisputably] a second layer to our world" [87],[7] have pointed out that humans in their interaction with AI, always interact with a model of the environment. Therefore, scheme 10, based on [7], visualizes a 'mutual perception relation', where humans and AI-technology change and shape each other. This scheme shows how the AI merges with a model ([ ]) that represents the environment – including humans – perceived by AI, based on data. At the same time, humans also have a perception of AI / the environment, which is formalized with the extra square brackets.

[20] The arrow indicates 'intentionality'.

Human – AI Relations	
5. Composite Intentionality	
Scheme	Human → (Technology → Environment)
Description	The human interacts with a by AI-mediated environment, similar to the hermeneutic relation. Hence, here the technology is not passively but actively engaging with the environment. Humans interact with an AI-mediated environment, similar to the hermeneutic relationship. However, the technology here is not passive, but actively engaged with the environment.
Example	Smart airport surveillance system
6. Cyborg relations	
Scheme	(Human / Technology) → Environment
Description	The human together with AI interact intentionally with the environment.
Example	Smart artificial hand for amputees
7. Immersion relation	
Scheme	Human ↔ Technology / Environment
Description	The AI merges with the environment and interacts with the human.
Example	Smart fridge
8. Relation of augmentation	
Scheme	(Human / Technology) → Environment & Human → (Technology - Environment)
Description	In augmented reality a cyborg relation is combined with a hermeneutic relation: i.e. The human with an AI head-mounted display has intentionality towards a by Technology interpreted environment
Example	Mixed Reality
9. Relegation	
Scheme	Human ← (Technology → Environment)
Description	The human intentionality has been taken over by AI, which means that the AI decides how the human interacts with the by AI-mediated environment.
Example	Commercial AI recommendation system in mixed reality [88].
10. Mutual perception relation	
Scheme	Human / [Human] ↔ [AI [Environment]] – Environment

Table 2

I amidst of AI

The crucial question here is: what is meant by the human self? In the context presented, this white paper distinguishes three aspects of what the self is: [21]

1. The Thinking Self
2. The Experiential Self
3. The Composite Self

According to (Post-)phenomenology, AI is part of the human self. The Thinking Self finds support in the conviction that the human mind can be computationally comprehended. This view is accompanied by the belief that what constitutes the mind can be seen independently of the body and can therefore be replaced by a computational substrate independent version. What AI may become as an artificial general or super intelligence can be seen as our Thinking Self. Moreover, AI is a derivative of human intelligence: 'a brain child'.
The Experiential Self is the self as it experiences the world without the intervention of thought. The Experiential Self also implies the disappearance of the separation between the self and the environment and the objectification of that what is experienced. The terms 'sensing' and 'feeling' fall under the provided description of experiencing. Emotions are classifications of these sensations, which imply an objectified view of oneself and fall therefore under the Thinking self.
The Composite Self, implies both the Thinking and Experiential Self, and is a Self conscious about both and therefore capable of relating to AI as a part of oneself.

[21] This distinction is rigorous and does not do justice to the fraught philosophical discussions surrounding this complex subject. This tripartite division serves the purpose of becoming aware of what we are as humans in order to remain critical of AI.

4 ***AI?*** A SERIOUS GAME

We have arrived at the embryonic outline of the mixed reality serious game ***AI?***, after describing AI's relation to the human and the human relationship with AI from both a computational-rational and phenomenological viewpoint, which exposed the problem of the abstraction of the human into measurable unities and the neglected importance of the human experience in human-AI collaboration.

What is ***AI?***

- ***AI?*** is an *interface* between the domain expert (professional) and an AI
- ***AI?*** *detects* other AI's and *exposes* the respective AI's intentionality and agency to the user
- ***AI?*** *teaches* computational thinking
- ***AI?*** *mentors* the user in becoming mindful
- ***AI?*** *deepens* the awareness of one's self as a thinking, experiential and composite being
- ***AI?*** *enables* a human-driven exploration of Human-AI alternatives
- ***AI?*** *stimulates* human critical reflection and creativity

AI? consists of four levels:

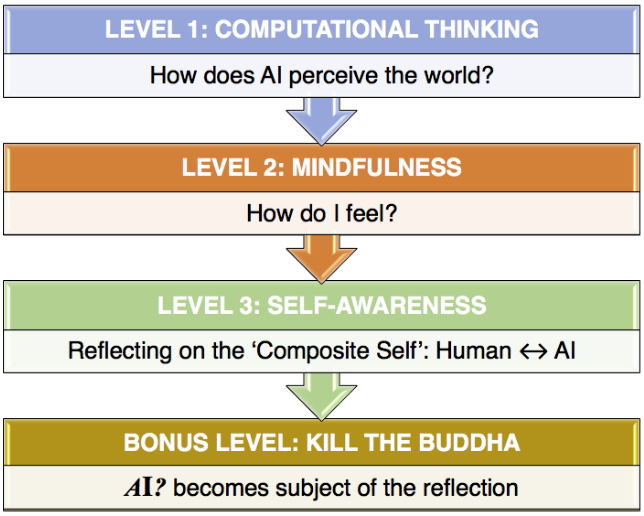


Figure 1

Level 1
The user must first become aware of the respective AI's understanding of the environment. By learning how AI sees the world partly differently from humans, the user can experience and understand that AI's intentionality and agency are not equal to human intentionality.

Level 2
Thereupon, the user will practice experiencing themselves as an experiential being.

Level 3
Next, the user explores how the respective AI is an extension of the user's intentionality or influences their intentionality or both. The Human-Technology / AI relation schemes will be the leading reference. This reflection enables the user to gain insight in the implications and consequences of this Human-AI collaboration. Based on these insights, the user can explore different alternatives and ultimately make human-beneficial decisions in their interplay with the respective AI. This gamified exploration embraces both computational thinking and putting experience into practice.

Bonus level
In the fourth level, the composite self-awareness obtained in level 3, should lead to a level 3 reflection on ***A*I*?***. The outcome of this reflection is decisive of what may become or be done with ***A*I*?***.

4.1 Serious Gaming

A description of serious gaming and serious gaming in mixed reality will now follow, in which a first attempt is made to shape the four levels on the basis of serious game guidelines.

Game-based learning
A game can be defined as “a system in which players engage in an artificial conflict, defined by rules, that results in a quantifiable outcome”[22]. Games have rules, enable player actions, provide responses, reflect previous actions, motivate the player to engage and are challenging [64]. A good game answers to several needs of players, such as knowing that it is possible to win, being challenged, and being able to ‘play God’[23].
Serious games[24] encourage active learning and user engagement through a game-like scenario [76],[53]. They are often a simulation of reality and allow people to train for unexpected situations without being put at risk [19],[76]. A successful serious game is both entertaining to play and achieves the learning goals [64].[25]
The following serious games design elements are used to accomplish the educational objectives [64]:

1) game mechanics
2) visual aesthetic, sound and narrative design
3) incentive system
4) Content and skills

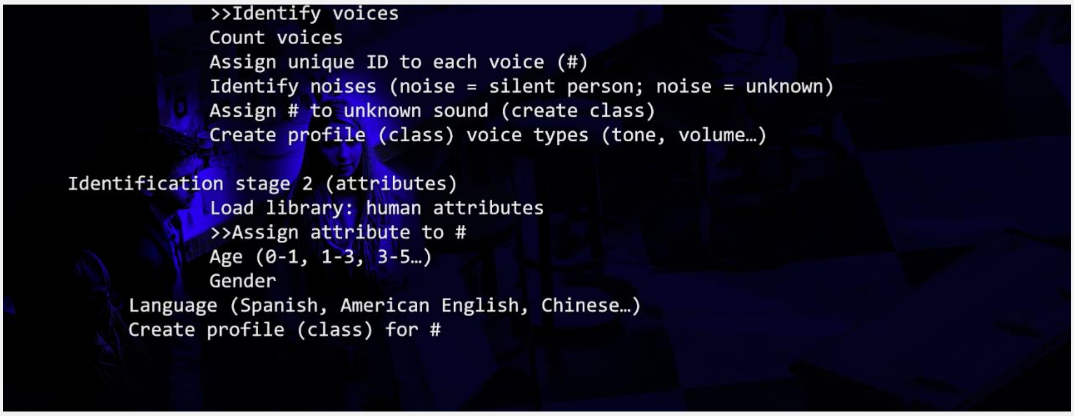
Game mechanics
Game mechanics help learners to acquire new knowledge or skills through interactions in the game world [62]. They are “the experiential building blocks of player interactivity [and are] repeated over and over throughout a game”[26]. Serious games can contain learning game mechanics, which define the essential learning interactions that take place and assessment game mechanics, which have a diagnostic goal [62],[64].[27]
The option to play God may potentially be a major factor in the successfulness of ***A*I*?***. One of the fears of professionals is the lack of personal agency and control when working with AI. ***A*I*?*** offers the possibility to view the situation in which the respective AI is active from a meta perspective. The learner knows the AI, themselves and is able to play with different scenario’s, which should include playing with unethical alternatives, such as destroying the AI or letting the AI do whatever it wants.[28]

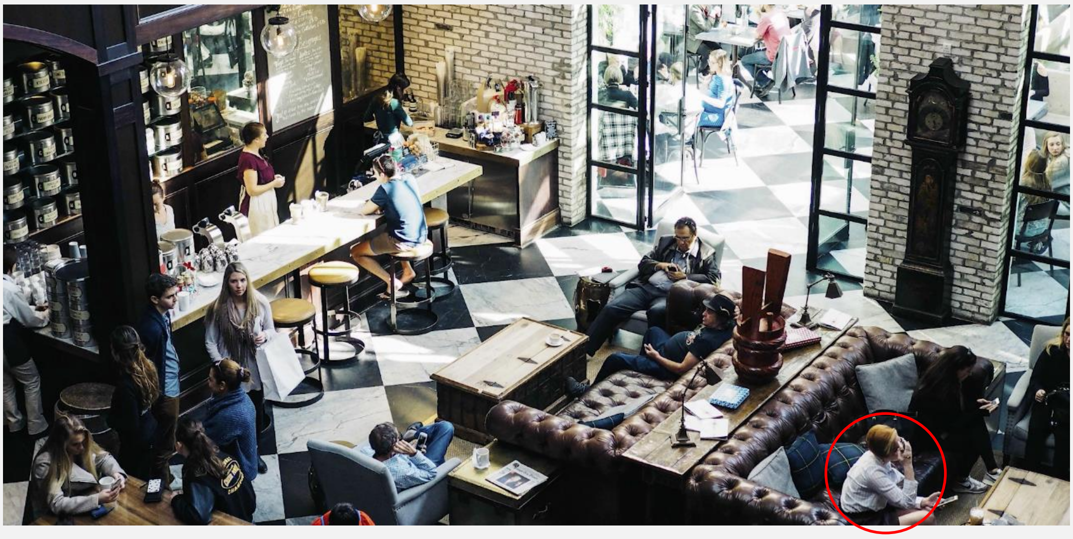
Visual aesthetic, sound and narrative design
An overview of visual game design for learning, applied to ***A*I*?***, can be found in the appendix, table 8.

[22] Salen and Zimmerman (2004) as cited in [64]: 5.
[23] “In virtual worlds, one can do almost anything. This exponential expansion of choice opportunities means that games can be a wellspring for autonomy satisfactions” in [70].
[24] The author has made no distinction in the literature between the terms 'serious gaming' and 'game-based learning'. Although the terms are used interchangeably, some literature uses different definitions.
[25] The increasingly popular term "gamification" differs from serious gaming, in that the former only adds game elements to reality to engage and encourage a person to behave in a certain way.
[26] Salen and Zimmerman (2004) as cited in [62]: 347.
[27] Table 7 in the appendix presents three game types that may be applicable to ***A*I*?***
[28] Having the option to do ‘wrong things’ explains the success of games like Grand Theft Auto [70]; and the failure of the BP energy serious game, where players were not allowed to put a power plant in the city, judging therefore the game as boring (source: author white paper, based on research for TNO (Netherlands Organization for Applied Scientific Research) (2013-2014)).

Level 1: Computational thinking
In regard to level 1, an artistic impression of how AI sees the world as has been provided by Silvio Carta (2021): “The Machine’s Eye” may serve as an design example.
The visualization shows how the AI screens the environment and draws conclusions based on data input. Humans are the result of a computational process. Everything is numerical. The representation of the physical is a set of lines of code. The learner reads the lines of code, and in doing so they imagine what the AI ‘sees’ and understand.

Screenshots of “The Machine’s Eye” by Silvio Carta





Level 2: Mindfulness
An overview of mindfulness and meditation applications, that may serve as design examples for level 2 can be found in the appendix (table 6).
Special attention in this context deserves the Mixed Reality (both VR and AR) mindfulness application HEALIUM**.** HEALIUM aims at creating self-awareness through bio- and neuro-feedback and creates a real-time visualization of one's current mental state. It may be fruitful to explore the possibilities of a similar application for ***AI?***.

Level 3: The Composite Self
The concept of 'digital twin' may be used for visualizations and experimentation of the alternative scenarios in level 3. A digital twin is an evolving digital profile of the historical and current behavior of a physical object or process to improve overall performance. Digital twins can be used for real-time testing in complex environments (even cities are being 'digitally twinned') without causing damage in the physical world. Current R&D explores the possibility of implementing the concept of 'digital twins' in mixed reality and in the metaverse.

Avatars and digital twins
AI? is intended to be used in a mixed reality. With regard to visual aesthetic design and narrative design elements, one can think of the development of the player's own avatar as a learning process. An avatar is a virtual agent representing someone's identity which are perceived by users as close representations of one's ideal self [81]. The potential of lifelike codec 2.0 avatars for ***AI?*** may also be explored.

Incentive system
Reward structures form the incentive system of a game. They provide feedback and direct the player's behavior [64]. There are two types of rewards: intrinsic rewards and extrinsic rewards. Intrinsic rewards relate to the learning objectives by playing the game. For example, a piece of knowledge that enables the learner to solve a problem or go to the next level [64],[81]. Extrinsic rewards, such as receiving points, are not directly related to the game play.

Potential intrinsic reward forms for ***AI?***[29]						
	Feedback messages	**Unlocking mechanisms**	**Developable avatars**	**Achievement systems**	**Game resources**	**Plot animations**
Level 1	✓	✓	✓	✓	✓	✓
Level 2	✓	✓	✓	✓	✓	
Level 3		✓	✓			✓
Bonus Level						✓

Table 3

Intrinsic motivations
Intrinsic motivations, such as exposing competence, autonomy and social connectedness are actions carried out for their own sake and are more successful in accomplishing education purposes [39]. Therefore, they appear to be more suitable for ***AI?***.

[29] This table presents a selection from a list provided by Wang and Sun (2011) as cited in [81].

Suggestions for intrinsic motivations to continue engaging with *AI?*	
Level 1	Gaining a sense of control by understanding what the respective AI is doing.
Level 2	Mastering tranquility. Although the main objective of ***AI?*** is not to be stress-relieving meditation application, it is important that the domain expert is able to feel relaxed and centered in their work environment.
Level 3	Curiosity to explore different scenarios and have the feeling to interact with AI from a personal agency. Since the goal of ***AI?*** is to empower the domain expert in his professional environment, it must be possible to apply the insights obtained in level 3 in the physical reality.
Bonus Level	Having full control over the destiny of ***AI?***. For example, the app can be completely removed, with the result that it can not be reinstalled. Or the learner may decide to share the scenarios from level 3 with a third party.

Table 4

Personalization of the game

Games can be personalized in several ways. One way is that the gamer can choose from several options to customize the game according to their preferences. The number of options can also be used as a reward system [33],[65],[70]. The other way is that the difficulty of the game is adjusted according to the player's performance to prevent the gamers from getting frustrated or bored [58],[40]. A table with adaptive variables applied to ***AI?*** can be found in the appendix (table 9).

In ***AI?*** the earlier mentioned avatar could, for example, be developed in level 1 and 2 and may be the reward to enter level 3. Here the avatar will be further developed by the learner in their exploration of different human-AI relations and future scenarios.

Content and Skills

Level 1: Computational thinking

AI? teaches computational thinking by applying it to the respective AI. It provides insight into the logic and reasoning of the AI and its objectives. It may show how it is connected with other AI's and how it perceives the domain expert.

AI? is an anthropomorphized explainable AI and does not provide necessarily insight into the actual datasets the AI works with and whether its inferences are correct.

Three study examples of the development of 'Computational Thinking Serious Games' can be found in the appendix.

Insight can be given into:

- How does the respective AI analyze data and make decisions?
- What are the goals?
- How can it reach them?
- How does the AI see the world the same and different?

Level 2: Mindfulness

The popular term 'mindfulness' is used for the deliberate practice of being in the present moment and becoming aware of sensations without thinking about these sensations. Mindfulness is not an inner mental observation, but just experiencing.

In the context of Phenomenological inquiry, [16] propose a mindfulness methodology based on Buddhist meditation techniques[30] for becoming experientially aware of one's past experiences. The process is described as "Suspension, Re-direction and Letting Go" and consists of three steps:

1. One takes time to decouple from one's activities both physically and mentally.
2. One turns from the external world towards one's interior subjective world.

[30] The two Buddhist mediation techniques mentioned are: Shamatha meditation, which is a concentration meditation; and Vipashyana meditation, i.e. an insight meditation: a continuous attention to bodily sensations and experiencing them without judgment.

3. One changes from active search to an accepting letting-arrive. One just sits and waits patiently without any effort to sensations to be revealed. There should also be no intention in what one is seeking.

The “Suspension, Re-direction and Letting Go” methodology appears to be useful in the context of ***A*I?** because it aims for re-experiencing past experiences.

The overview of mindfulness and meditation applications that can be found in the appendix focus on stress reduction, better sleep and improving focus, among others, through breathing exercises, mental body scans and shifting from focus to awareness. The popular meditation application 'Headspace' uses gamification to engage its users. There are also applications that train conscious breathing by visualizing breathing. The Moodflow app is not a meditation app, but asks the user to keep track of how he is feeling. Mindfulness is not concerned with labeling the feelings that are experienced. However, further research is needed to understand how mindfulness achievement can be verified and used as input to Level 3.[31]

Level 3: The Composite Self
In this level, the Human-AI relation will be explored on the basis of (a selection of) the ten Post-phenomenological Human-Technology / AI relations. The reflection on these relations should result in a composite (Human + AI) understanding of one’s ‘self’ applied to the respective AI. Level 3 aims at guiding the learner into a novel understanding of reality from where new possibilities rise.
The learner becomes aware of (1) the interaction that the human does not exist independent from AI, and that they mutually shape each other; (2) that AI has a model of the world (including the user) based on what can be measured (ideally a scientific understanding) and does not necessarily resembles the reality as an individual sees it; (3) that the learner themselves have a perception of the reality, that is partly formed by one’s experiences.
The next step will be to explore different scenario’s with the respective AI and to experience and reflect on the consequences of these different scenario’s.

Bonus Level: Kill the Buddha
This level assesses whether the learner is able to critically reflect on ***A*I?**.

4.2 Serious Gaming in Mixed Reality

Mixed reality
This white paper anticipates a professional future where mixed reality is the order of the day. There is currently no consensus on what exactly 'mixed reality' is. What is understood by it, is based on developments in augmented reality (AR) and virtual reality (VR). Some experts argue that in ten years AR, VR and mixed reality will exist just as one form. Mixed reality has been described as a reality-virtuality continuum [56], meaning that mixed reality is a reality with a variable number of virtual elements, but not fully virtual. This reality-virtuality continuum may be shared by many users or experienced by one [75].
AR adds virtual (digital) dynamic and context specific objects and information to the real world [88], which can be sensed through a handheld device, smart glasses, holographic display or head-mounted display.[32] The information can be visual, audible or provided as sensory feedback.[33]

[31] “Meditative traditions, for example, also give instructions for evaluating and validating achieved results, by demanding the careful verification of the presence and properties of the results arrived at along the way of their own paths of experience” [16]: 78.
[32] Currently, there are two types of AR: marker-based and marker-less. Marker-based AR, is AR that adds an overlay to a clearly defined physical objects. Marker-less AR projects virtual 3D objects on a designated spot in a physical environment.
[33] The AR AI does this by using sensors such as digital compasses, a thermal positioning sensor and an accelerometer, to read data from each physical place to predict the focused area where the user is navigating.

VR is the full immersion of users in a computer technology created digital interactive three-dimensional world in which the objects have a sense of spatial presence.[34] Currently, VR is used for training purposes and is expected to become popular as the metaverse, which is believed to be "a persistent digital world that is inhabited by digital twins of people, places, and things"[35].

Mixed reality in professional environments
Mixed reality is increasingly used in both the professional and consumer markets, with a dominant increase in entertainment, healthcare, engineering and defense. [36],[37] AR provides real-time support for complex tasks. AR is developed and deployed to optimize the performance of professionals and the efficiency of internal processes, such as: improving productivity, quality control, accuracy, compliance. VR is mainly being used for training purposes.

Medical specialist care will be discussed as an example for ***AI?*** in a moment. However, it is worth noting that the industrial application of AR is numerous and consists of training, asset identification, remote knowledge transfer, expert assistant, process design, virtual manuals, complex manufacturing, among others. Therefore, the manufacturing or process industry may be also worth exploring as a test case for AI?.

Serious Gaming in Mixed Reality
In a mixed reality game-based learning environment, both the personal and environmental contexts are sensed and supported with data from databases, so that the mixed reality can respond to the learner's feedback [33]. The added value of mixed reality in a game-based learning environment is that it "mediates and enhances visual and cognitive perceptions"[76], enables a learning experience independent of a specific physical environment [23] and can be activated by location or object (image recognition) [53].
In regard to mixed reality learning, the following three design principles are applicable [23]:

1. Enable and then challenge
2. Drive by gamified story
3. See the unseen, i.e. making visible what is invisible.[38]

In the case ***AI?*** all three design principles may apply	
Level 1	See the unseen.
Level 2	?
Level 3	Drive by gamified story.
Bonus Level	Enable and then challenge. For example, disabling the application to invoke a 'Composite Self' response.

Table 5

Ubiquitous learning and computational thinking
Learning in mixed reality can also be referred to as ubiquitous learning [50]. This implies "an everyday learning environment that is supported by mobile and embedded computers and wireless networks in our everyday life"[39], which is in fact an 'AI powered Internet of Things' (AIoT).
The authors point out that learning 'computational thinking in AIoT through AR' makes students more aware of its complexity. This article provides a learning table, which, among other things, teaches students to recognize patterns observed through the AR device. This concept could be used in ***AI?*** to learn to see the world the way AI does.

[34]This description is based on the definition of VR as provided by NASA.
[35] Such as Microsoft's Mesh for Teams.
[36]https://www.reportlinker.com/p06246486/Augmented-Reality-Services-Global-Market-Report.html?utm_source=GNW
[37] https://lumusvision.com/augmented-reality-trends-infographic/
[38] This form is often used in scientific learning environments [23].
[39] Ogata et al. 2009 as cited in [46]: 3361

5 *AI?* APPLIED: THE MEDICAL SPECIALIST

The text below provides a futuristic vision of the application of AI and mixed reality to support the medical specialist, illustrating the degree of immersion of the domain expert in AI technology, followed by an exploratory impression of how *AI?* can be used.

AI in medicine
AI is becoming an indispensable part of medicine, motivated by developing more effective and efficient healthcare, which should lead to a reduction in healthcare costs [61]. In medicine, AI is increasingly used for data collection and analysis, diagnostics[40], clinical decision support and improved vision [38].[41] AI in combination with vision and large data sets supports the specialist in recognizing deviations in patterns. AI will be able to detect malignant tumors at an early stage, recognize other diseases and predict disease risk more accurately than its human counterpart. The combination of patient-specific data with relevant large amounts of data results in patient-tailored treatment.

AI in the operating room of the future
In the operating room (OR) of the future, all devices will be connected and form a cyber-physical-human system. This will allow the AI to model the process of surgical work and provide the right information at the right time in this highly dynamic and complex environment. AI will provide data-driven decision support, context-aware assistance, cognitive robotic assistance and workflow analysis support. AI will support the surgeon with the precision removal of tumors by means of advanced visioning [38].[42] AI robots will take over physical tasks, such as suturing[43] and needle insertion [9]. These robots will have besides vision, also cognitive intentionality and agency [60]. In AI robotic precision surgery, the surgeon only needs to supervise the operation. Although the robots are able to act autonomously, they will only act on command [9].

Mixed reality in the OR
Augmented and virtual reality combined with AI will be used for medical diagnostics, routine screening, patient record viewing, telemedicine, remote patient monitoring and therapies and lifelong surgery training [13]. In the OR, mixed reality will support the surgeon by presenting interactive 3D anatomical models of the patient, along with medical illustrations and patient history. Mixed reality will simplify preoperative planning and surgeons will use mixed reality to visualize and discuss their approaches with other specialists in different geographic locations, before and during surgery [48],[72]. Mixed reality will provide additional expert support, reducing the need for technicians in the OR [83]. Surgeons will use mixed reality for complex procedures during surgery, where AI-devices are cross-linked with high-precision optical navigators, which “effectively reduces the surgeon's dependence on spatial experience and imagination” [72],[93:1]; and as haptic technology (artificial touch) advances, surgeons will also be able to perform operations remotely with the support of mixed reality AI.

Trust in AI in medicine
The majority of physicians are known to have strong reservations about automating medical tasks and strongly oppose autonomous AI in surgery [37]. In general, ethical discussions argue that AI in medicine should not make autonomous decisions.[44]
A doctor may ask an AI decision support system what the most impactful intervention for preventing hospital readmission for a patient is [59], to which the AI will provide a data-based answer. Explainable and transparent AI will provide insight into how it came to this conclusion and the physician will make a rational decision whether or not to trust the AI.[45] However, the

[40] Currently, mainly in radiology, pathology, dermatology and cardiology.
[41] https://futuretodayinstitute.com/mu_uploads/2022/03/FTI_Tech_Trends_2022_Book06.pdf, https://www.ibm.com/watson-health/solutions/
[42] https://news.lenovo.com/ai-analysis-revolutionizes-early-stage-tumor-removal-techsomed/, https://www.mevis.fraunhofer.de/en/solutionpages/liver-surgery.html.
[43] ‘Smart Tissue Autonomous Robot’ outperformed surgeons on the surgery of a pig intestine [60].
[44] That said, it is recognized that AI does influence decision making [59] and that the physician may fall victim to automation bias or loose their skills, because they do not have to use them.
[45] [59] argues for a normative attitude towards trust in medical AI. Normative trust is related to rational trust and is based on predetermined norms agreed upon by both parties [78]. This allows the doctor to rationally analyze the situation and decide on the best option with the least uncertainty.

doctor may not understand all the steps involved in the AI's decision-making process, or have a ‘gut feeling’, based on past experiences, that may be contrary to the AI's recommendation. In this case, this means that emotional trust is lacking.[46]
It is likely that in the future, humans will not be able to oversee all the processes and conclusions on which the AI has based its recommendation. This event will leave the healthcare professional with seemingly no choice but to trust blindly. However, [47] presented an alternative view of emotional trust, stating that trust in technology should be seen as an act of confidence. Here, the relationship with AI is confidently taken as a starting point to reflecting on the human relations with it [87].

Reflecting on the trust-relation with AI
This white paper has defended that it is crucial for the reflection on this inextricable relationship, that one becomes aware of how both AI and oneself experience the environment. This composite self-awareness should ultimately lead to a fruitful collaboration with AI on a personal level, as well as enable the professional to formulate answers to questions that concern the entire organization regarding the use and implications of AI [61],[59]
Translated back to the specialist and AI decision support system, a pressing issue that would benefit from the latter's constructive support is the bias in diagnostic algorithms due to erroneous or incomplete data due to human bias AI [61]. When the domain expert understands how AI reasons and what the human part is in AI reasoning, they may be more willing to work with AI developers to improve the algorithms to achieve the desired outcome.

AI? *applied*

Level 1: Computational Thinking[47]

Fictional example

1. As a medical specialist prepares for surgery using a mixed reality head-mounted display, ***AI?*** detects multiple AIs in the operating room.
2. The specialist, motivated to evaluate their work and to center themselves, decides to engage within a time frame of a week with ***AI?***.
3. ***AI?*** asks the surgeon to choose from the detected AI’s, providing them an overview (this overview of AI’s could be visualized in different ways).
4. The surgeon decides to explore their relation with the AI suturing robot.
5. ***AI?*** provides insight in the computational processes of the robot, resulting in the actions as performed.
   e.g. It distinguishes a different color or structure of the skin to decide where to suture. Since ***AI?*** anticipates a network of AI devices, ***AI?*** will show how the robot is connected and how it interacts with other AI’s. It will also show how the robot perceives the human: e.g. by means of their voice commands or a prior action performed with another AI device to which the robot is connected. ***AI?*** will provide insight into the robot by asking questions to the surgeon or by having them perform the task from the robot's point of view.
6. If the physician has proven to have understood the suturing robot, they will gain access to level 2.

Level 2: Mindfulness
In this level, the specialist will take the time to center themselves experientially. As proposed by [16], this can be an exercise of bringing back the surgeon's experience in the operating room, withdrawn from any judgment.
A HEALIUM-like virtual environment may be provided. The surgeon may decide if they want their brain functioning to be monitored or that they enter this mindful state without the use of

[46] It is unlikely that the option will exist to not engage with the decision-support AI. It is also assumed here, that the physician will be responsible for the decision.
[47] While some defend that more research is needed to support the intuition that computer science reasoning should be incorporated into the medical curriculum [57], nevertheless computational thinking is already beginning to gain momentum in medical school today.

the head-mounted display. The specialist may also choose a specific environment to become mindful.

It requires further research in how and if the gained experiential awareness should be measured, or if the surgeon provides the feedback that they are ready themselves, based on their own insight (which would be like a phenomenological psychological interview). All in all, level 2 is about one's personal experience, and not an objectively measured state to achieve. That being said, it may also be optional to engage in the measuring of bio or neural feedback.

Level 3: The Composite Self

1. The specialist explores which Post-phenomenological relation with the AI suturing robot is appropriate. The understanding of these schemes and the exploration should be provided in an attractive game form.
2. The specialist concludes that it is a cyborg relation ((Human / Technology) → Environment) inside the mutual perception relation (Human / [Human] ↔ [AI [Environment]] – Environment), since the robot also actively interacts with the surgeon.
3. The specialist may then compare their perception of the procedure, based on both their experience and cognition, with the computational model the robot has of the situation. In this phase, overlap and discrepancy may be discovered.

This relationship may be visualized in the form of a digital twin with a user-customized avatar or a computationally created avatar. In this digital twin environment, the specialist may explore alone or together with the robot, what the differences would be if the specialist had done the suturing themselves, or how the robot will do the suturing in a surgery that is scheduled for the next day. Because the specialist understands now how the suturing robot relates to other AI devices, as well as the role the surgeon themselves plays in this cyber-physical-human system, they may make suggestions on how another device can be connected to improve the desired outcome.

AI may impair some of the physician's skills because they don't have to use them. A level 3 reflection, may bring this to light. It may be that in the future only the robot is allowed to suture, because it is more accurate, less painful, etc. – similar to the devices used today, which cannot be replaced by devices from 50 years ago. The experiential awareness gained in level 2 should have led to acceptance. Yet, the reflection can, for example, also lead to the insight that the surgeon misses the precision work with the hands, and they realize that it is not important that it is related to medicine. This insight also improves the specialist's sense of self.

As input for improvements at the supervisory level, the level 3 reflection can also shed personal light on other systemic issues, such as regarding liabilities, insurance, malpractice, etc.

Level 4: Kill the Buddha

Level 4 is a reflection on the further utility of ***A*I*?***. Only the physician can decide if the guidance of ***A*I*?*** is still useful or interferes with their personal agency. No data will be extracted from ***A*I*?*** without consent, and only the user may give permission to share, for example, a simulation from Level 3. This level is an exercise in critical awareness, leaving ***A*I*?***'s future in the hands of the user.

The integration of ***A*I*?*** *in the work environment*

An important practical question is, when will a domain expert engage with ***A*I*?***. Doctors, but also other professionals in highly demanding environments, are uttermost occupied. ***A*I*?*** should therefore be seen as a formation: a formation in Human-AI collaboration.[48] This means

[48] Although there are positive results in the application of serious gaming in medical education, serious gaming is not yet widely applied. A systematic review of serious games in medical education has not shown that they are more effective than traditional medical education. However, the majority of the serious games are to be used in a class room setting [35], while against the background of lifelong learning, ***A*I*?*** is not meant as a serious game in the class

that dedicated time should be scheduled to engage with ***A*I?**. Since ***A*I?** is presented as an integrated in mixed reality application that detects other AI's, ***A*I?** will be functional during the professional interaction with other AI. A surgeon will not have time to reflect upon what they are doing when preparing and performing the surgery. ***A*I?** will detect and save the involvement with the respective AI, to come back to it at a later time when it is appropriate for the physician.

room, but as a self-improvement tool in the professional setting. In addition, a research paper on a serious game that prepares medical students to work in the emergency room, by offering a virtual emergency room with features found in state-of-the art computer games, has shown to be effectively in teaching nontechnical skills [12].

6 CONCLUDING THOUGHTS

This white paper has attempted to highlight a problem in the collaboration between humans and AI, namely the consequences of reducing humans to their reason and understanding humans and their agency in terms of measurability; and has provided a first step for an integrated solution in the form of a mixed reality serious game that trains humans in self-awareness for human critical reflection and creativity.

This white paper has proposed to extend R&D in Human-AI collaboration with the subjective experience of the involved professional, and defends that only then, human beneficial AI collaboration can take place. ***AI?*** is a first proposition how this may be done.

Much research is focused on human-centric AI with an emphasis on transparency and explainability to gain user trust and improve collaboration. The scientific approach to humans makes it easier for AI to achieve its goals involving humans. The anthropomorphization of AI's Human-Machine Interface focuses on human emotions, but does not make AI intrinsically human. The paper has sought to challenge common thinking about human-AI collaboration and highlights the importance of human value as an experiential being in relation to AI to sustain a hybrid organization aligned with human insights and values.

A broader application of ***AI?***

At the dawn of the era of AI and mixed reality, where an increasing amount of actions is permanently carried out by AI without human intervention and where humans are being integrated into the AIoT via smartphones and other connected devices, it is justifiable to familiarize humans with the foregoing and make them aware of how AI technologies influence their intentionality. ***AI?*** can therefore also be developed for the non-professional who wishes to become aware of their interactions with AI, instead of uncritically trusting the AI. 'Critical' should not be understood as taking an opposite attitude or as intellectual problem-solving, but as awareness, both intellectually and experientially.

Based on current R&D, one can posit that Brain-Computer Interfaces (BCI) will be used by the general population in the not too distant future. A BCI measures and digitizes human brain activity, whereupon the brain can directly interact with AI devices. In that case, it will be highly desirable to have a sense of self and not to be controlled by the BCI, as depicted in science fiction. ***AI?*** can be linked to the BCI to support the person in becoming aware of the AI, their experiential self and the purpose of the collaboration, which can be adjusted if necessary.

Development challenges and further research

This white paper leaves the technical reader with many unanswered questions. Questions that may have arisen, are: 'how will ***AI?*** detect other AI and retrieve insight in its agency? And, 'will ***AI?*** be trusted by other AI?' ***AI?*** is intended as a self-awareness application and should not become an AI detection tool. ***AI?*** gathers data, connects to other AI, but is not allowed to intentionally collect data of the learner and share it with third parties.

One of the challenges with regard to game development, is that ***AI?*** strives to go beyond measurability and mentors the learner to become their own critical yardstick, while at the same time ***AI?*** is an AI application, which can only exist on the basis of quantification.

This white paper has provided an applied philosophical argument for the importance to develop an application like ***AI?***. Therefore, it should be read as an invitation to AI scientists, serious game developers, and AI-automated organizations who acknowledge the importance of a truly human organization to further explore the practical feasibility of ***AI?***.

ACKNOWLEDGEMENTS

I am grateful for the constructive and motivational conversations, feedback and input of Emile van Bergen, Florian Güldner, Wyn Jones, Nana Efua Lawson, Igor Rust and Eileen McGinnis; and in particular my dear friend Prof. Andrea Aparo von Flüe, who allowed me to express my intuitions and helped me with his critical questions and suggestions to concretize them in meaningful sentences.

The images of the artwork “The Machine’s Eye” are used with the permission of the author of the artwork Dr. Silvio Carta.

The artwork 'Human-AI collaboration' on the title page was generated by artificial intelligence, due to a lack of personal inspiration and available human resources. The author's intentional keywords in the UI were "Human AI Collaboration". Development time: less than a minute...
Nightcafe Creator (https://creator.nightcafe.studio)) 2022, July 19

Parts of this white paper are based on the author's ongoing PhD research (Department of Philosophy, University of Twente, The Netherlands) on 'Human Critical and Creative Thinking in the Era of Artificial Super Intelligence’.

This white paper is supported in whole by ‘The Notre Dame-IBM Tech Ethics Lab’. Such support does not constitute endorsement by the sponsor of the views expressed in this publication.

ABOUT THE AUTHOR

Ida Helena Rust is an expert advisor concerned with the optimization of Human-AI systems collaboration and a business developer for science & technology, acting as an interface between STEM professionals and people with a non-technological background.
She has an academic background in film dramaturgy (BA, MA) and philosophy (BA) and is currently studying for her PhD Philosophy, the influence of Artificial Super Intelligence (ASI) as an advanced AI of Everything on human critical and creative thinking. She explores this question by using Postphenomenological theory and comparing ASI with three concepts of God: the personal God, the pantheistic God, and the process God. The question underlying her research is: "what can we call 'human' in the era of ASI?".
Her motivation and personal philosophy are the preservation of human individuality in smart technological environments to strengthen the resilience of both the human and the cyber-physical-human system as a process.

REFERENCES

[1] T. Araujo, N. Helberger, S. Kruikemeier, and C.H. de Vreese, "In AI we trust? Perceptions about automated decision-making by artificial intelligence," *AI & Soc.,* vol. 35, pp. 611–623, 2020, doi: 10.1007/s00146-019-00931-w.

[2] C. Aydin, "The Posthuman as Hollow Idol: A Nietzschean Critique of Human Enhancement," *The Journal of medicine and philosophy.,* vol. 42, no. 3, pp. 304-327, 2017, doi:10.1093/jmp/jhx002.

[3] C. Aydin, M. González Woge, and P.P. Verbeek, "Technological Environmentality: Conceptualizing Technology as a Mediating Milieu". *Philos. Technol.* Vol. 32**,** pp. 321–338, 2019, doi: 10.1007/s13347-018-0309-3.

[4] C. Aydin, *"Extimate Technology Self-Formation in a Technological World,"* New York and London, Routledge, 2021.

[5] C. R. B. Azevedo, K. Raizer and R. Souza, "A vision for human-machine mutual understanding, trust establishment, and collaboration," *2017 IEEE Conference on Cognitive and Computational Aspects of Situation Management (CogSIMA)*, 2017, pp. 1-3, doi: 10.1109/COGSIMA.2017.7929606.

[6] A. Baier, "Trust and Antitrust," *Ethics*, vol. 96, no. 2, pp. 231-260, Jan., 1986. Available: https://www.jstor.org/stable/2381376.

[7] J.J. Benjamin, A. Berger, N. Merrill, and J. Pierce, "Machine Learning Uncertainty as a Design Material: A Post-Phenomenological Inquiry," in *Proceedings of the 2021 CHI Conference on Human Factors in Computing Systems (CHI '21). Association for Computing Machinery, New York, NY, USA,* Article 171, pp.1–14, 2021, doi: 10.1145/3411764.3445481.

[8] A.F. Blackwell, L. Church, and T.R.G. Green, "The abstract is 'an enemy': Alternative perspectives to computational thinking," *in Proceedings PPIG'08, 20th annual workshop of the Psychology of Programming Interest Group*, pp. 34-43, 2008.

[9] S. Bodenstedt, M. Wagner, B.P. Müller-Stich, J. Weitz, S. Speidel: "Artificial Intelligence-Assisted Surgery: Potential and Challenges," *Visc Med.,* vol. 36, pp. 450-455, Dec. 2020, doi: 10.1159/000511351.

[10] N. Bostrom, "Superintelligence Paths, Dangers, Strategies," Oxford, UK: University Press, 2014.

[11] D.J. Chalmers, "The Singularity: A Philosophical Analysis," *Journal of Consciousness Studies*, vol. 17, no. 9-10, pp. 7-65, January 2010.

[12] S.H. Chon, F. Timmermann, T. Dratsch, N. Schuelper, P. Plum, F. Berlth, R.R. Datta, C. Schramm, S. Haneder, M.R. Späth, M. Dübbers, J. Kleinert, T. Raupach, C. Bruns, and Kleinert, "Serious Games in Surgical Medical Education: A Virtual Emergency Department as a Tool for Teaching Clinical Reasoning to Medical Students," *JMIR Serious Games*, vol. 7, no. 1, March 2019, doi: 10.2196/13028.

[13] Condino, S.; Gesi, M.; Viglialoro, R.M.; Carbone, M.; Turini, G. Serious Games and Mixed Reality Applications for Healthcare. Appl. Sci. 2022, 12, 3644, doi: 10.3390/app12073644.

[14] K.E. Culley and P. Madhavan, "A note of caution regarding anthropomorphism in HCI agents," *Computers in Human Behavior,* vol. 29*,* no. 3, pp. 577–579, 2013, doi: 10.1016/j.chb.2012.11.023.

[15] S. Darwall, "Trust as a Second-Personal Attitude (of the Heart)," in *The Philosophy of Trust*, P. Faulkner and T. Simpson, ed. Oxford, UK: Oxford University Press, 2017.

[16] N. Depraz, F. Varela, and P. Vermersch, "On Becoming Aware A pragmatics of experiencing," Amsterdam: John Benjamins Publishing Co., 2003.

[17] J. Dewey, "How we think," Sep. 2011. [E-book no. 37423]. Available: https://www.gutenberg.org/ebooks/37423.

[18] J. Dewey, "Logic, The Theory of Inquiry," New York, NY, US: Henry Holt and Company, 1938.

[19] D. Djaouti, J. Alvarez, J.P. Jessel, and O. Rampnoux, "Origins of Serious Games," in *Serious Games and Edutainment Applications,* M. Ma, A. Oikonomou, and L. Jain, eds. London: Springer, 2011, doi: 10.1007/978-1-4471-2161-9_3.

[20] P. Domingos, "The Master Algorithm," Basic Books, 2015. [E-book].

[21] J. Domenicucci and R. Holton, "Trust as a Two-Place Relation," in *The Philosophy of Trust*, P. Faulkner and T. Simpson, ed. Oxford, UK: Oxford University Press, 2017.

[22] H.L. Dreyfus, "Why Heideggerian AI failed and how fixing it would require making it more Heideggerian," *Philosophical Psychology*, vol. 20, no. 2, pp. 247-268, 2007, doi: 10.1080/09515080701239510.

[23] M. Dunleavy, "Design Principles for Augmented Reality Learning," in *TechTrends,* vol. 58, pp. 28–34, 2014, doi: 10.1007/s11528-013-0717-2.

[24] L. Eliot, "AI Ethics Is Especially Vexed By That AI Confinement Problem, Including The Knotty Particulars For Confining Autonomous Self-Driving Cars," in: Forbes website, May 2022. Available: https://www.forbes.com/sites/lanceeliot/2022/05/05/ai-ethics-is-especially-vexed-by-that-ai-confinement-problem-including-the-knotty-particulars-for-confining-autonomous-self-driving-cars/?sh=2213a6a3500f; [Accessed May 07, 2022].

[25] D.C. Elton, "Self-explaining AI as an Alternative to Interpretable AI," in: Artificial General Intelligence. AGI 2020, *Lecture Notes in Computer Science*, vol. 12177, B. Goertzel, A. Panov, A. Potapov, and R. Yampolskiy, R., eds., Springer, Cham., doi: 10.1007/978-3-030-52152-3_10.

[26] P. Faulkner and T. Simpson, "Introduction," in *The Philosophy of Trust*, P. Faulkner and T. Simpson, ed. Oxford, UK: Oxford University Press, 2017.

[27] A. Ferrario, M. Loi, and E. Viganò, "In AI We Trust Incrementally: a Multi-layer Model of Trust to Analyze Human-Artificial Intelligence Interactions," *Philos. Technol.* Vol. 33, pp. 523–539, 2020, doi: 10.1007/s13347-019-00378-3.

[28] M. Franssen, G.-J. Lokhorst, and I. v.d. Poel, "Philosophy of Technology", *The Stanford Encyclopedia of Philosophy*, 2018, E.N. Zalta, eds. Available: https://plato.stanford.edu/archives/fall2018/entries/technology/. [accessed: July 25, 2022].

[29] S. Gallagher, "What Is Phenomenology?," in: *Phenomenology*. *Palgrave Philosophy Today*, London, UK: Palgrave Macmillan, 2012, doi: 10.1057/9781137283801.

[30] S. Gallagher and D. Zahavi, "Phenomenological Approaches to Self-Consciousness", in *The Stanford Encyclopedia of Philosophy,* E. N. Zalta, ed., 2021. Available: https://plato.stanford.edu/archives/spr2021/entries/self-consciousness-phenomenological/. [accessed: July 25, 2022].
[31] D. Gambetta, "Can We Trust Trust?", in *Trust: Making and Breaking Cooperative Relations*, D. Gambetta, ed., Oxford: University of Oxford, pp. 213-237, 2000. [E-book]. Available: http://www.sociology.ox.ac.uk/papers/ gambetta213-237.pdf.
[32] S.J. Gershman, E. J. Horvitz, and J. B. Tenenbaum, "Computational rationality: A converging paradigm for intelligence in brains, minds, and machines," *SCIENCE***,** Vol. 349, no. 6245, pp. 273-278**,** Jul. 2015, doi: 10.1126/science.aac6076.
[33] F. Giannakas, G. Kambourakis, A. Papasalouros, and S. Gritzalis, "A critical review of 13 years of mobile game-based learning," *Education Technology Research and Development,* vol. 66, pp. 341–384, 2018, doi: 10.1007/s11423-017-9552-z.
[34] E. Glikson and A. W. Woolley, "Human Trust in Artificial Intelligence: Review of Empirical Research," *Academy of Management Annals,* vol. 14, no. 2, Aug. 2020, doi: 10.5465/annals.2018.0057.
[35] I. Gorbanev, S. Agudelo-Londoño, R.A. González, A. Cortes, A. Pomares, V. Delgadillo, F.J. Yepes, Ó. Muñoz, "A systematic review of serious games in medical education: quality of evidence and pedagogical strategy," *Medical Education Online*, vol. 23, no. 1, art. 1438718, Dec. 2018, doi: 10.1080/10872981.2018.1438718.
[36] R. van Gulick, "Consciousness", *The Stanford Encyclopedia of Philosophy,* E. N. Zalta, ed., 2021. Available: https://plato.stanford.edu/archives/win2021/entries/consciousness/. [accessed: July 26, 2022].
[37] A.A. Gumbs, I. Frigerio, G. Spolverato, R. Croner, A. Illanes, E. Chouillard, and E. Elyan, "Artificial Intelligence Surgery: How Do We Get to Autonomous Actions in Surgery?," *Sensors,* Basel, Switzerland, vol. 21, no. 16, art. 5526, 2021, doi: 10.3390/s21165526.
[38] D.A. Hashimoto, G. Rosman, D. Rus, and O. R. Meireles, "Artificial Intelligence in Surgery: Promises and Perils," Annals of surgery, vol. 268, no. 1, pp. 70–76, 2018, doi: 10.1097/SLA.0000000000002693.
[39] B.D. Homer, C. Raffaele, and H. Henderson, "Games as Playful Learning: Implications of Developmental Theory for Game-Based Learning," in *Handbook of Game-based learning*, J.L. Plass, R.E. Mayer, and B.D. Homer, eds., Cambridge, Massachusetts, US: The MIT Press, 2019, pp. 25-52, 2019.
[40] D. Hooshyar, M. Pedaste, Y. Yang, L. Malva, G.-J. Hwang, M. Wang, H. Lim, H., and D. Delev, "From Gaming to Computational Thinking: An Adaptive Educational Computer Game-Based Learning Approach," *Journal of Educational Computing Research*, vol. 59, no. 3, pp. 383–409, 2021.
[41] J. Hughes, "The Deskilling of Teaching and the Case for Intelligent Tutoring Systems", *J. Eth. Emerg. Tech.*, vol. 31, no. 2, pp. 1–16, Dec. 2021.
[42] D. Ihde, "Technology and the Lifeworld: From Garden to Earth," Bloomington, Indiana, US: Indiana University Press, 1990.
[43] D. Ihde, L. Malafouris, "Homo faber Revisited: Postphenomenology and Material Engagement Theory," *Philos. Technol.*, vol. 32, pp. 195–214, 2019, doi: 10.1007/s13347-018-0321-7.
[44] K. Jones, "Trust as an Affective Attitude," in *Ethics*, vol. 107, no. 1, pp. 4-25, Oct., 1996. Available: http://www.jstor.org/stable/2382241. [Accessed: July 08, 2009]
[45] A. Kakar, "A Theory of Agile Software Development," *SAIS 2020 Proceedings 32*, Nov. 2020. Available: https://aisel.aisnet.org/sais2020/32.
[46] Kinshuk and S. Graf, "Ubiquitous Learning," in: *Encyclopedia of the Sciences of Learning*, N.M. Seel, ed., Boston, MA, US: Springer, 2012, doi: 10.1007/978-1-4419-1428-6_224
[47] A. H. Kiran, P.-P. Verbeek, "Trusting Our Selves to Technology," *Know Techn Pol,* vol. 23, pp. 409–427, 2010, doi: 10.1007/s12130-010-9123-7.
[48] R. P. Kumar, E. Pelanis, R. Bugge, H. Brun, R. Palomar, D. L. Aghayan, Å. A. Fretland, B. Edwin, and O. J. Elle, "Use of mixed reality for surgery planning: Assessment and development workflow," in *Journal of Biomedical Informatics*, vol. 112, art. 100077, 2020. [Supplement]. Doi: 10.1016/j.yjbinx.2020.100077.
[49] J. C. Lester, R. D. Spain, J. P. Rowe, and B. W. Mott, "Instructional Support, Feedback, and Coaching in Game-Based Learning," in *Handbook of Game-based learning,* J.L. Plass, .E. Mayer, and B.D. Homer, eds., Cambridge, Massachusetts, US: The MIT Press, 2019, pp. 209-239, 2019.
[50] N. Liberati, "Augmented reality and ubiquitous computing: the hidden potentialities of augmented reality," in *AI & Soc*, vol. 31, pp. 17–28, 2016, doi: 10.1007/s00146-014-0543-x2014.
[51] Y.S. Lin, S.Y. Chen, C.W. Tsai, and Y.H. Lai, "Exploring Computational Thinking Skills Training Through Augmented Reality and AIoT Learning," *Frontiers in psychology*, vol. 12, art. 640115, doi: 10.3389/fpsyg.2021.640115.
[52] S. Lockey, N. Gillespie, D. Holm, and I. A. Someh: "A Review of Trust in Artificial Intelligence: Challenges, Vulnerabilities and Future Directions," in *Proceedings of the 54th Hawaii International Conference on System Sciences*, Jan. 2021, doi: 10.24251/HICSS.2021.664.
[53] M.M. Marques and L. Pombo, "Game-Based Mobile Learning with Augmented Reality: Are Teachers Ready to Adopt It?," in: *Project and Design Literacy as Cornerstones of Smart Education. Smart Innovation, Systems and Technologies*, M. Rehm, J. Saldien, and S. Manca, eds., vol. 158, Singapore: Springer, 2020, doi: 10.1007/978-981-13-9652-6_19.
[54] J. McCarthy, "Ascribing mental qualities to machines," Stanford, CA. US: Stanford University, Computer Science Department, 1979. Available: http://www-formal.stanford.edu/jmc/. [Accessed: July 26, 2022].
[55] R. McIntyre and D. W. Smith, "Theory of Intentionality," in *Husserl's Phenomenology: A Textbook,* J. N. Mohanty and W. R. McKenna, eds., Washington, D. C., US: Center for Advanced Research in Phenomenology and University Press of America, 1989, pp. 147-79.
[56] P. Milgram and F. Kishino, "A taxonomy of mixed reality visual displays," in *IEICE Transactions on Information Systems,* vol. E77-D, No.12, Dec. 1994. Available: https://cs.gmu.edu/~zduric/cs499/Readings/r76JBo-Milgram_IEICE_1994.pdf.
[57] P. Musaeus, D. Tatar, and M. Rosen, *Medical Computational Thinking: Computer Scientific Reasoning in the Medical Curriculum,"* in: *Emerging Research, Practice, and Policy on Computational Thinking.*

Educational Communications and Technology: Issues and Innovations, P. Rich, and C. Hodges, eds. Springer, Cham. 2017, doi: 10.1007/978-3-319-52691-1_6.

[58] B. Nelson and Y. Kim, “Multimedia Design Principles in Game-Based Learning,” in *Handbook of Game-based learning,* J.L. Plass, R.E. Mayer, and B.D. Homer, eds., Cambridge, Massachusetts, US: The MIT Press, 2019, pp. 307-329, 2019.

[59] P.J. Nickel, “Trust in medical artificial intelligence: a discretionary account,” *Ethics and Information Technology*, vol. 24, no. 7, 2022. https://doi.org/10.1007/s10676-022-09630-5.

[60] S. Panesar, Y. Cagle, D. Chander, J. Morey, J. Fernandez-Miranda, and M. Kliot, “Artificial Intelligence and the Future of Surgical Robotics,” Annals of Surgery, vol. 270, no. 2, pp. 223-226, Aug. 2019, doi: 10.1097/SLA.0000000000003262.

[61] S.Y. Park, P. Kuo, A. Barbarin, E. Kaziunas, A. Chow, K. Singh, L. Wilcox, and W. S. Lasecki, “Identifying Challenges and Opportunities in Human-AI Collaboration in Healthcare,” in Conference Companion Publication of the *2019 on Computer Supported Cooperative Work and Social Computing (CSCW '19)*, Association for Computing Machinery, New York, NY, USA, pp. 506–510. https://doi.org/10.1145/3311957.3359433.

[62] S. Pawar, F. Tam, and J. L. Plass, “Emerging Design Factors in Game-Based Learning: Emotional Design, Musical Score, and Game Mechanics Design,” in *Handbook of Game-based learning,* J.L. Plass, R.E. Mayer, and B.D. Homer, eds., Cambridge, Massachusetts, US: The MIT Press, 2019, pp. 347-367, 2019.

[63] P. Phillips, C. Hahn, P. Fontana, A. Yates, K. Greene, D. Broniatowski, and M. Przybocki, “Four Principles of Explainable Artificial Intelligence,” in *NIST Interagency / Internal Report (NISTIR), National Institute of Standards and* Technology, Gaithersburg, MD, US, 2021, doi: 10.6028/NIST.IR.8312. Available: https://tsapps.nist.gov/publication/get_pdf.cfm?pub_id=933399. [Accessed July 26, 2022].

[64] J.L. Plass, B. D. Homer, R. E. Mayer, and C. K. Kinzer, “Theoretical Foundations of Game-Based and Playful Learning,” in *Handbook of Game-based learning,* J.L. Plass, R.E. Mayer, and B.D. Homer, eds., Cambridge, Massachusetts, US: The MIT Press, 2019, pp. 3-25, 2019.

[65] J.L. Plass and S. Pawar, “Adaptivity and Personalization in Game-Based Learning,” in *Handbook of Game-based learning,* J.L. Plass, R.E. Mayer, and B.D. Homer, eds., Cambridge, Massachusetts, US: The MIT Press, 2019, pp. 263-283, 2019.

[66] K.F. Pfaffinger, J.A.M. Reif, E. Spieß, and R. Berger, “Anxiety in a digitalised work environment,” *Gr Interakt Org,* vol. 51, pp. 25–35, 2020, doi: 10.1007/s11612-020-00502-4.

[67] D. Romero, P. Bernus, O. Noran, J. Stahre, and Å. Fast-Berglund, “The Operator 4.0: Human Cyber-Physical Systems & Adaptive Automation Towards Human-Automation Symbiosis Work Systems,” in: ,*et al. Advances in Production Management Systems. Initiatives for a Sustainable World. APMS 2016. IFIP Advances in Information and Communication Technology,* vol. 488, Springer, Cham., 2016, doi: 10.1007/978-3-319-51133-7_80.

[68] D. Rudrauf, A. Lutz, D. Cosmelli, J.-P. Lachaux, and M. le van Quyen, “From autopoiesis to neurophenomenology: Francisco Varela's exploration of the biophysics of being,” in *Biol. Res.* [online]. Vol. 36, no. 1, pp. 27-65, 2003, doi: 10.4067/S0716-97602003000100005.

[69] S.J. Russell and P. Norvig, “Artificial Intelligence, A Modern Approach,” Harlow, Essex, England: Pearson Education Limited, 2016.

[70] M. Ryan and C.S. Rigby, “Motivational Foundations of Game-Based Learning” in *Handbook of Game-based learning,* J.L. Plass, R.E. Mayer, and B.D. Homer, eds., Cambridge, Massachusetts, US: The MIT Press, 2019, pp. 153-177, 2019.

[71] M. Ryan, “In AI We Trust: Ethics, Artificial Intelligence, and Reliability,” in *Sci Eng Ethics,* vol. 26, pp. 2749-2767, 2020, doi: 10.1007/s11948-020-00228-y.

[72] J. A. Sánchez-Margallo, A. Juan, M.C. de Plaza, R.A. Fernández Anzules, and F.M. Sánchez-Margallo, “Application of Mixed Reality in Medical Training and Surgical Planning Focused on Minimally Invasive Surgery,” *Frontiers in Virtual Reality*, vol. 2, art. 692641, 2021, doi: 10.3389/frvir.2021.692641. Available: https://www.frontiersin.org/article/10.3389/frvir.2021.692641.

[73] J. Smith, "Self-Consciousness", *The Stanford Encyclopedia of Philosophy* Edward N. Zalta, ed., 2020. Available: https://plato.stanford.edu/archives/sum2020/entries/self-consciousness/. [Accessed July 26, 2022].

[74] J.M. Smith, “Is Computational Thinking Critical Thinking?,” in: , *et al. Expanding Global Horizons Through Technology Enhanced Language Learning. Lecture Notes in Educational Technology,* Springer, Singapore, 2021, doi: 10.1007/978-981-15-7579-2_11.

[75] M. Speicher, B.D. Hall, and M. Nebeling, “What is Mixed Reality?,” in CHI Conference on Human Factors in *Computing Systems Proceedings (CHI 2019), May 4–9, 2019, Glasgow, Scotland, UK. ACM, New York, NY, USA.* Doi: 10.1145/3290605.3300767.

[76] L. Ştefan and F. Moldoveanu, “Game-based Learning with Augmented Reality – From Technology’s Affordances to Game Design and Educational Scenarios,” in *The 9th International Scientific Conference eLearning and software for Education Bucharest,* pp. 105-114, April 25-26, 2013*,* doi: 10.12753/2066 026X-13-124.

[77] O. Stone and D. Zahavi, “Phenomenology and Mindfulness,” in *Journal of Consciousness Studies*, vol. 28, no. 3-4, pp.158-185, 2021. [Abstract].

[78] M. Taddeo, “Defining Trust and E-trust: Old Theories and New Problems,” in *International Journal of Technology and Human Interaction (IJTHI) Official Publication of the Information Resources Management Association,* vol 5, no. 2, pp. 23-35, 2009.

[79] M. Taddeo, “Trust in Technology: A Distinctive and a Problematic Relation,” in *Know. Techn. Pol*., vol. 23, pp. 283–286, 2010, doi: 10.1007/s12130-010-9113-9.

[80] M. Taddeo and L. Floridi, “The case for e-trust,” in *Ethics Inf. Technol.,* vol. 13, pp. 1-3, 2011, doi: 10.1007/s10676-010-9263-1.

[81] F. Tam and S. Pawar, “Emerging Design Factors in Game-Based Learning: Incentives, Social Presence, and Identity Design,” in *Handbook of Game-based learning,* J.L. Plass, R.E. Mayer, and B.D. Homer, eds., Cambridge, Massachusetts, US: The MIT Press, 2019, pp. 367-387, 2019.

[82] F.J. Varela, E. Thompson, and E. Rosch, “The embodied mind,” Cambridge, Massachusetts, US: The MIT

Press, 2016 [E-book].
[83] R. Veloso, R. Magalhães, A. Marques, P. Gomes, J. Pereira, “Mixed Reality in an Operating Room Using Hololens 2—The Use of the Remote Assistance from Manufacturers Technicians during the Surgeries,” in Eng. Proc., vol. 7, no. 54, 2021, doi: 10.3390/engproc2021007054.
[84] J. W. Velthuijsen, W. van Tol, and A. Hagen, “Human Value in the Digital Age,” in *PricewaterhouseCoopers*, Dec. 2017, Available: https://www.pwc.nl/nl/assets/documents/pwc-human-value-in-the-digital-age.pdf.
[85] P.-P. Verbeek, “Wat things do,” Pennsylvania, US: The Pennsylvania State University Press, 2005
[86] P.-P. Verbeek, “Cyborg intentionality: Rethinking the phenomenology of human–technology relations,” *Phenom Cogn Sci,* vol. 7, pp. 387–395, 2008, doi: 10.1007/s11097-008-9099-x.
[87] P.-P. Verbeek, “Designing the Public Sphere: Information Technologies and the Politics of Mediation,” in *The Onlife Manifesto, Being Human in a Hyperconnected Era*, L. Floridi, ed., Springer Open, pp. 217-229, 2015.
[88] G. Wellner, “Postphenomenology of augmented reality,” in *Relating to Things*, *Design, Technology and the Artificial,* H. Wiltse, ed., London, UK: Bloomsbury Visual Arts, pp.173-189, 2020.
[89] P. Werkhoven, M. Neerincx, L. Kester, “Telling autonomous systems what to do,” in *European Conference on Cognitive Ergonomics 2018 Proceedings*, *8 pages*. Doi: 10.1145/3232078.3232238.
[90] J.M. Wing, “Computational Thinking,” in *Communications of the ACM*, vol. 49, no. 3, pp. 33-35, March 2006, doi: 10.1145/1118178.1118215.
[91] J.M. Wing, “Computational Thinking: What and Why?,” 2010. Available: https://www.cs.cmu.edu/link/research-notebook-computational-thinking-what-and-why. [Accessed: July 26, 2022].
[92] World Economic Forum, “The future of jobs, Employment, Skills and Workforce Strategy for the Fourth Industrial Revolution,” Jan. 2016. Available: https://www3.weforum.org/docs/WEF_Future_of_Jobs.pdf. [Accessed: July 26, 2022].
[93] R. Yang, C. Li, P. Tu, A. Ahmed, T. Ji, and X. Chen, “Development and Application of Digital Maxillofacial Surgery System Based on Mixed Reality Technology,” in *Front. Surg.* vol. 8, art. 719985, 2022, doi: 10.3389/fsurg.2021.719985
[94] D. Zahavi, “Phenomenology of self,” in *The Self in Neuroscience and Psychiatry*, T. Kircher and A. David, Eds. Cambridge: Cambridge University Press, 2003, pp. 56–75.
[95] D. Zahavi, “Internalism, Externalism, and Transcendental Idealism,” in *Synthese* , vol. 160, no. 3, pp. 355–374, 2008.
[96] D. Zahavi, “Phenomenology of reflection: Section III, chapter 2, Universal structures of pure consciousness," *in Commentary on Husserl's "Ideas I",* A. Staiti, ed., Berlin, München, Boston: De Gruyter, 2015, pp. 177-194, Doi: 10.1515/9783110429091-011
[97] D. Zahavi, “Phenomenology The Basics,” Abingdon, Oxon, UK: Routledge, 2019
[98] D. Zhang, N. Maslej, E. Brynjolfsson, J. Etchemendy, T. Lyons, J. Manyika, H. Ngo, J. C. Niebles, M. Sellitto, E. Sakhaee, Y. Shoham, J. Clark, and Raymond Perrault, “The AI Index 2022 Annual Report,” Palo Alto, CA, US: Stanford Institute for Human-Centered AI, Stanford University, March 2022. Available: https://aiindex.stanford.edu/report/.

APPENDIX: SUPPLEMENTARY INFORMATION

Three study examples for the development of Computational Thinking Serious Games:

1) D. Hooshyar, M. Pedaste, Y. Yang, L. Malva, G.-J. Hwang, M. Wang, H. Lim, H., and D. Delev, "From Gaming to Computational Thinking: An Adaptive Educational Computer Game-Based Learning Approach," *Journal of Educational Computing Research*, vol. 59, no. 3, pp. 383–409, 2021.

2) A. Malizia, D. Fogli, F. Danesi, T. Turchi and D. Bell, "TAPASPlay: A game-based learning approach to foster computation thinking skills," *2017 IEEE Symposium on Visual Languages and Human-Centric Computing (VL/HCC)*, 2017, pp. 345-346, doi: 10.1109/VLHCC.2017.8103502.

3) S. Werneburg, S. Manske, and H.U. Hoppe, "ctGameStudio – A Game-Based Learning Environment to Foster Computational Thinking," *in ICCE 2018 - 26th International Conference on Computers in Education, Main Conference Proceedings*, pp. 543-552, 2018.

An overview of mindfulness and meditation applications, that may serve as design examples for level 2:

Breathe+ (https://apps.apple.com/us/app/breathe-simple-breath-trainer/id1106998959)

Breathe+ is an application that visualizes your breathing in a simple way. This app is designed to help you learn how to breathe "properly" for full benefits from meditative breathing, including being able to hold your breath for up to 30 seconds and inhaling or exhaling for even short periods of time.

Headspace (https://www.headspace.com/headspace-meditation-app)

Mindfulness and meditation application

Calm (https://www.calm.com)

Calm features hundreds of calming exercises, helpful breathing techniques, and sleep stories narrated by celebrities.

Insight Timer (https://insighttimer.com)

Insight Timer is a free meditation application with guided meditations, sleep music tracks and talks led by the top meditation and mindfulness experts, neuroscientists, psychologists and teachers from renowned universities.

The Mindfulness App (https://themindfulnessapp.com)

The Mindfulness App has over 300 guided meditations and courses from experts around the world.
The app also offers personalized meditation options, reminders to keep you mindful throughout the day, and statistics to track in your meditation journal.

Mindfulness with Petit BamBou (https://www.petitbambou.com/en/)

This application features many guided meditation courses from 3 to 50 minutes long in multiple languages, such as English, Spanish, and German. The app has courses based on the science of positive psychology, cognitive behavioral therapy, and mindfulness.

Moodflow (https://www.moodflow.co)

Moodflow allows the user to note and track your emotions, thoughts, moods, and mental wellbeing. It comes with an easy-to-use interface with options to customize the app; you can set a background video or image, choose your preferred language, set a passcode lock, and change the calendar's appearance.

Ten Percent Happier *(https://www.tenpercent.com)*

The Ten Percent Happier app offers, e.g.:

- Video and meditation combos that make it fun and straightforward to learn to meditate
- A library of 500+ guided meditations (including Vippasana)
- Bite-size stories, wisdom, and inspiration that you can listen to while on the go
- Quick meditations that fit into your busy life

UCLA Mindful (https://www.uclahealth.org/marc/ucla-mindful-app)
This app offers, e.g.: • Basic Meditations in both English and Spanish • Wellness Meditations for people suffering from challenging health conditions • Informative videos exploring how to get started, supportive meditation postures, and the science of mindfulness • A timer to meditate on your own

Table 6

The following game types can be used for assessment: puzzle-like, situated assessments and discovery based learning environments [58].

Game types			
	Puzzle-like	**Situated assessments**	**Discovery based learning environments (2D / 3D)**
Level 1	✓	✓	✓
Level 2		✓	
Level 3		✓	✓
Bonus Level			✓

Table 7

An overview of elements of game design for learning has been provided by [58].

Elements of game design for learning: an initial suggestion for AI?	
Level 1	• Incentive systems (points) • Musical score (correct and incorrect sounds) • Teach new knowledge and skills (learn circuit)
Level 2	
Level 3	• Visual aesthetic design (e.g. avatar, in-world objects, virtual worlds) • Narrative design • Practice and reinforce existing knowledge and skills
Bonus Level	• Visual aesthetic design (e.g. avatar, in-world objects, virtual worlds) • Narrative design • Practice and reinforce existing knowledge and skills

Table 8

Games have different variables upon which they can adapt, categorized as cognitive, motivational, affective and sociocultural variables. In the case of ***AI?*** variables such as current knowledge, ability to self-regulate, persistence, emotional state, self-perception could be thought of as relevant to use. These variables can be used to assess the user's progression, thereupon adapting the learning environment [58].
Table 9 is an estimation of relevant variables and subject to further inquiry.

Adaptive Variables				
	Cognitive	**Motivational**	**Affective**	**Sociocultural**
Level 1	• Current knowledge • Developmental level	• Individual interest • Situational interest • Goal orientation • Self efficacy • Persistence		
Level 2		• Individual interest • Situational interest • Goal orientation • Persistence	• Emotional state	
Level 3	• Current knowledge • Developmental level • Cognitive abilities / skills	• Individual interest • Situational interest • Goal orientation • Self efficacy • Persistence	• Emotional state • Emotion regulation • Attitudes	• Social context • Cultural context • Identity/self-perception • Social agency
Bonus Level		• Individual interest • Situational interest • Goal orientation • Self efficacy	• Emotion regulation • Attitudes	• Identity/self-perception • Social agency

Table 9

Contact: i.r.h.rust@utwente.nl