On Children's Exploration, Aha! Moments and Explanations in Model Building for Self-Regulated Problem-Solving

Vicky Charisi*[1], Natalia Díaz-Rodríguez*[2], Barbara Mawhin[3] and Luis Merino[4]

[1]Joint Research Centre, European Commission, Seville, Spain.
[2]DaSCI Andalusian Institute in Data Science and Computational Intelligence, Univers. of Granada, Spain
[3]Human Factors Department, EBT-Salient Aero Foundation, Spain
[4]Service Robotics Laboratory, University Pablo de Olavide, Seville, Spain

Abstract

In certain problem-solving tasks that require Human-AI interactions, a mutual understanding of the reasoning behind the performed actions can benefit both humans and artificial agents. However, identifying and predicting the cognitive strategies involved in such a hybrid setting, especially in novel, self-regulated exploratory tasks, is a challenging endeavour. Our aim is to identify behavioural properties relevant to young children's cognitive strategies that are present in problem-solving, with an emphasis on the Aha! moment as an intermediate step between exploratory actions, that typically relate to the development of tacit knowledge, and the generation of explanations that requires explicit knowledge. We use data from existing, previously published, behavioural studies with children 5 to 7 years old to explore these mechanisms in two self-regulated problem-solving tasks. In addition, we reflect on our observations of an Artificial Agent (Q-learning algorithm) that learns to solve the same task. Our findings indicate that while in current reinforcement learning practice, detecting the moment of the cognitive transformation of the problem representation normally translates into observing convergence curves of the objective functions being optimized, in young children this involves more complex behavioural properties, such as verbal metacognition. These behavioural processes can be used as a proxy for the identification of the Aha! moment. Finally, we propose a conceptual map which integrates the observed behaviours that are used to detect, communicate and corroborate learning both in humans and machines and we discuss the association of children's exploratory behaviours, the Aha! moments and ultimately their explanation generation.

Keywords

Explainability, Child development, Human intelligence, Problem-solving, Behavioural indicators, Explainable AI

1. Introduction

For effective hybrid environments where humans collaborate with Artificial Intelligence (AI) systems to make a decision, a mutual understanding of the reasoning behind certain actions or recommendations can be of catalytic importance.

Explainability is one of the features that supports mutual understanding and trust development [1], and can be considered as an interface through which machine learning models can be explained towards a customized and diverse set of audiences [2], debugged, and audited. For the generation of explanations, though, implicit knowledge should become explicit, which often includes the cognitive process known as the Aha! moment or insight. We adopt the definition of the Aha! moment in problem solving as a sudden transformation of the problem representation [3, 4]; this differs from the solution retrieval reached by an analytical, multistep strategy, through which the solver searches long-term memory for potential algorithms, mental schemas, analogies or factual knowledge.

In this paper we seek to clarify what behavioural manifestations indicate the occurrence of the Aha! moment in children performing certain problem-solving tasks and instantiate a conceptual map of strategies which are used to detect, communicate and corroborate learning both in humans and machines. The ultimate goal is providing a richer test-bed of procedural protocols and tests to more broadly assess learning in machines, beyond a single metric or loss optimization.

1.1. Inspiration by children's problem-solving

Reverse engineering human intelligence can usefully inform AI and machine learning. The exploration of fundamental cognitive processes that can be informative for AI approaches often requires focusing on infants or young children in the context of structured or unstructured activities [5, 6, 7]. Self-regulated play, for example, that allows children to perform exploratory actions and come up with insights and discoveries in problems they generated has previously been correlated with the de-

EBeM'22: IJCAI-ECAI Workshop on AI Evaluation Beyond Metrics, July 25, 2022, Vienna, Austria.
(*) Equal contribution.
vasiliki.charisi@ec.europa.eu (V. Charisi*)
0000-0001-7677-027X (V. Charisi*); 0000-0003-3362-9326 (N. Díaz-Rodríguez*); 0000-0003-2857-4922 (B. Mawhin); 0000-0003-4927-8647 (L. Merino)

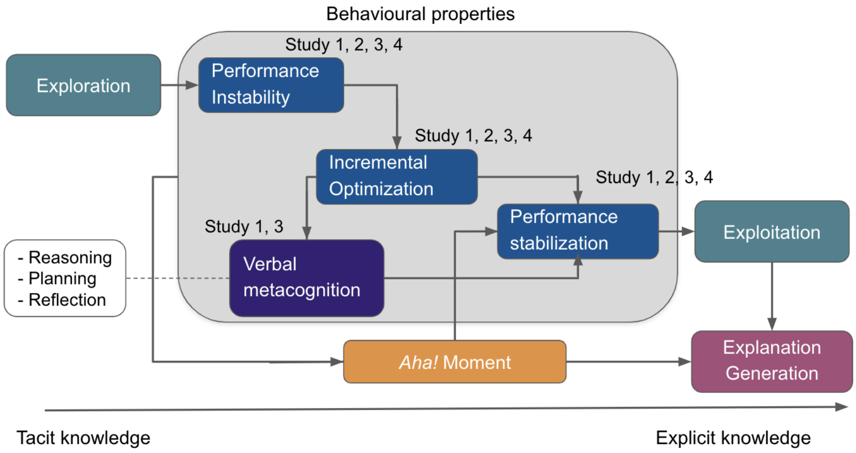

Figure 1: Behavioural properties as a proxy for the identification of the Aha! moment in children's problem solving process. The proposed conceptual map includes behaviours for the evaluation of the transition from tacit knowledge (appearing in the phase of exploration) toward explicit knowledge (appearing in the phase of exploitation). The properties include non-verbal behaviours and verbal metacognitive manifestations (reasoning, planning and reflection). The Aha! moment appears as part of the transition from tacit to explicit knowledge and functions as an indicator for the generation of explanations.

velopment of their implicit knowledge and their gradual understanding of the surrounding physical world [8]. However, what cognitive process are mobilized for the transformation of tacit into explicit knowledge in young children? And what behavioural properties can be used as a proxy for the identification of those processes?

Based on a series of behavioural studies with children 5 to 7 years of age, we identify behavioural properties relevant to cognitive processes that are present in problem-solving tasks, with an emphasis on identifying the Aha! moments, as an intermediate step between exploratory actions and the generation of explanations (see Fig. 1), aiming to inform current and future approaches on explainable AI (XAI).

2. Relevant Work

2.1. Problem-Solving in Young Children

To understand the fundamentals of problem-solving as a cognitive process, developmental psychologists have extensively explored the involved faculties and the ways they interact with each other. To this end, classic and contemporary work has examined various tasks that were used depending on the child's age and areas of interest. Bruner, for example, laid out a plan for the development of skilled action [9]. First there is intention, then an assembling of "constituent acts". They initially occur out of order but later become properly sequenced to reach the goal. Bruner emphasised the role of exploratory behaviour and play prior to achieving skilled action. Flexibility and higher order acts become possible through reorganization of component acts and modularization. Although Bruner's examples came from infants in the first year of life, his ideas have been applied to the acquisition of more complex skills beyond infancy. Additionally, he argued that play is the best way to promote development as it can occur with any physical material or with imagination, alone or with others and can take place in various settings [10]. The connection of play with the development of fundamental cognitive processes and human learning has been well-established [11, 12, 13]. Self-directed and intrinsically motivated goal generation and problem-solving are among children's cognitive tools that affect their overall development [7, 14]. In free play, children set novel goals, discover unexpected information, and invent problems they would not otherwise encounter. In this context, children apply exploratory processes that allow them to progressively reduce uncertainty about their environment [14].

In this context, a problem is defined as a situation in which a solver needs to change a given state to a desired one but there are obstacles. There are different types of problems such as the routine problem vs. the non-routine problem. The first one refers to a situation in which the solver knows a solution method whereas the second is when the solver has to create a solution method. There is also the well-defined problem where the state, goal and set of operators are clearly defined. It is opposed to the ill-defined problem where the elements are not clearly defined.

The problem solving process occurs when a person has to invent a way to solve it following two main stages: the problem representation and the problem solution. The solvers need to comprehend the problem and create a model of the problem situation. Then, they have to build a solution by using processes of planning, executing and they have to monitor it using awareness and control. It implies cognitive and metacognitive processes. Problem solving is always domain-specific but the thinking by analogy strategy seems to be almost always successful. Thinking of a related problem already known and even better, already solved, helps for success. An application of this is the heuristics which allow a solver to go faster to an acceptable solution even if it is not perfectly accurate. Considering the bounded rationality of humans, heuristics allows us to make judgements, choices and adapt our behaviours efficiently. This is closely related to the concepts of "social learning" and "adaptation" in human development.

In order to solve a problem, two mental representations are needed: one of the current state and one of the goal state. As it is goal-oriented and contextualized, a plan detailing the solution step by step is required. A constant monitoring process is also required as each move has consequences that can bring the solver closer to or further to the desired goal state. It also requires mental flexibility and thus, inhibitory control [15]. If a first chosen solution seems to be inappropriate, the solver has to adapt his strategy.

2.2. Social learning

Social learning is a crucial component of human intelligence, allowing us to rapidly adapt to new scenarios, learn new tasks, and communicate knowledge that can be built on by others [16]. The work of Lev Vygotsky who put forward this view already in the 1920s takes into account factors such as the language development and cultural influences in the cognitive development of children [17]. From his perspective, mental functioning and development rely on an interdependence between individual and social processes. When learners, whatever their age, participate in joint activities, they gain new abilities and strategies to better understand the world and adapt to it. This process is also mediated by signs and tools such as language and mnemonic techniques. Vygotsky folds them in the category of semiotics means. They are considered as a cornerstone for knowledge co-construction and can help independent problem-solving activity. This leads to the difference of what a learner can do with or without help as he described under the concept of the Zone of Proximal Development (ZPD). The social interaction with the use of linguistic and cultural tools facilitate the internalization of knowledge and its transformation into cognitive tools supporting the development of new cognitive functions. The latter aspect has been considered for the design of artificial agents that are able to interact with others and internalize these interactions in a similar way as humans [18, 19].

The process named scaffolding is described as a process that enables a child or novice to solve a task or achieve a goal that would be beyond his unassisted efforts [20]. To achieve more complex tasks (like problem-solving), it is necessary to combine simpler skills in order to achieve a higher level of competence. This promotes cognitive growth. The shared space of an activity involving collaboration mechanisms between peers is also at great importance whether it is a human or an artificial agent [21, 22].

2.3. Insight in Problem-Solving

Most commonly, this phenomenon is called the "Aha!" experience describing the moment when a person gets the solution to a problem that up to this point had left her puzzled. In cognitive science it is referred to as insight problem solving and it is accompanied by a feeling of satisfaction for the solver. It has been related to creative thinking [23, 24] and includes an exploratory phase where divergent thinking takes place, especially during the early stages of the problem solving process. This allows the person to produce new ideas or connect existing ideas. The second phase is the convergent thinking phase where a solution should be elected by synthesizing, analyzing and monitoring the matching degree of the current result to the expected one. Although the experience of insight is sudden and can seem disconnected from the immediately preceding thought, recent research shows that insight is the culmination of a series of brain states and processes operating at different time scales. Elucidation of these precursors suggests interventional opportunities for the facilitation of insight [3], including concurrent verbalization [25]. As for every problem solving, most of these strategies rely on a constant restructuring of the mental representation of the problem. One way in which explicit knowledge manifests itself is through the formation of causal inferences and the generation of explanations that, in research with children, are used for detecting gaps in their causal knowledge.

2.4. Explanation generation

Regarding explanation generation, there is a large body of works in various fields. But it always implies an explainer and an explainee with their own respective characteristics. Of particular interest across the fields is the role of the Theory of Mind ie. the ability of a person to attribute mental states to the consequent behaviours of herself or others [26]. The selection and evaluation processes of explanations depend on the explainer and explainee, but also on the characteristics of the context. The nature of an interaction for explaining is different in kindergarten between the teacher and a young child than between the cockpit desk and the pilots during a flight. The role of beliefs has also been raised recently as a cornerstone. An explanation does not necessarily needs to be consistent with a person's beliefs but should help promoting a revision component [27] thus allowing the evolution of the internal representations. Human explanations from social sciences became an integrated part of Artificial Intelligence (AI) through the XAI field in order to provide explanatory agents and to facilitate interactions between humans and machines.

3. Methodological approach

We aimed to identify behavioural properties in young children's problem-solving process that have been allocated on the transition from tacit knowledge to the development of explicit knowledge and the generation of explanations. We used two types of problem-solving activities, an open-ended task (computer-supported music-making) and the cognitive task of the Tower of Hanoi (ToH). We considered the above-mentioned theories and we conducted three behavioural studies to explore children's processes in various settings. In addition, w to solve the same problem of the ToH. For the purpose of this paper, we take a case-study methodological approach. Case studies are in-depth investigations of a single person, group, event or community which are approached from a qualitative perspective [28]. All the included case-studies meet the following criteria: (i) The sample consists of children aged 5 to 7 years old, and (ii) the setting facilitates children's self-regulated activities. In order for us to ensure the necessary variability among the case studies, we included (i) open vs. non open-ended tasks and (ii) different types of social contexts (collaboration with two children, collaboration of one child with a robot, hybrid collaboration with two children and a robot, see Figure 5).

It should be noted that any comparison among the studies was outside the scope of this paper; rather, our goal is to make a synthesis of the results as appeared in different settings. For this reason, we only provide the necessary overall findings for each study and we adopt a qualitative reflective approach for one representative case-study per experiment. The selection of the case studies was based on their relevance to the purpose of this work and on their representativeness of children's average behaviour in specific settings.

4. Empirical Studies: A Selection of Use Cases

This section presents a line of empirical evidence that have contributed to our identification of behavioural indicators that facilitated the transition from tacit to explicti knowledge. For each study we first describe the original goal of the study, analysis of the data that are relevant to the current work and the corresponding findings and we reflect on how this contributes to the purposes of this work.

Figure 2: Self-regulated music-making setting: a. The Reactable, a table-top interface for sound synthesis and the touchscreen version of it with two participants; b. The Sibelius Groovy, a music-making software for children and the setting with two children.

4.1. Study 1: Identification of behaviours

The scope of the study was to identify behaviours that emerge spontaneously when children are involved in an open-ended problem solving activity and to observe their development over time. As such, the setting of the study was based on ethnographic methodological principles and there was no adult intervention during the activities. Open-ended tasks without adult intervention provide the space for children to pursue their goals in self-regulated and intrinsically motivated manner. We designed a naturalistic behavioural study in a school-setting with $N = 16$ young children (5-6 years old) who were invited to compose music in pairs with the use of two dedicated screen-based software packages in a weekly basis over a period of maximum 8 weeks (Fig. 5.) The children were only asked to create music with the sounds provided by the digital tool. No other intervention was performed from the experimenter. The observations included $1795.51 min$ of video data which were transcribed based on an annotation scheme with a taxonomy of behaviours in relation to children's cognitive process, social interactions and affective engagement. For the purposes of this paper we only focus on the first category. A detailed description of the study appears in [29].

4.1.1. Data analysis

For the elaboration of the data we used the approach of microgenetic analysis [30]. The microgenetic method is defined by three properties: (a) observations span a period of rapid change in competency; (b) the density of

Code	Behaviour	Occurrence (%)
C1	Spontaneous musicking	11.05
C2	Sound exploration	15.87
C3	Assessment	27.38
C4	Reasoning	18.6
C5	Deliberate musicking	13.06
C6	Planning	14.04

Table 1
The taxonomy of behaviours that emerged during the open-ended self-regulated task of children's collaborative music-making and the percentage of occurrence per behaviour.

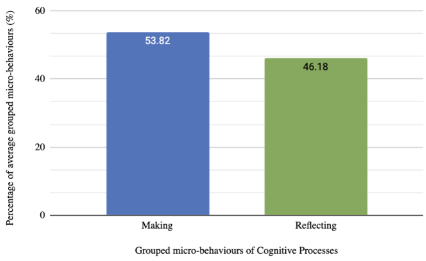

Figure 3: Average percentage of children's behaviours in Study 1, Making (C1, C2, C5 and C6) and Reflecting (C3 and C4).

observations is high relative to the rate of change; and (c) the observations are subjected to an intensive, trial-trial analysis to infer the processes that give rise to change. The annotation of the data was based on children's verbal and non-verbal behaviours and the corpus included 7063 annotated behaviours. The taxonomy of the behaviours that related to children's cognitive processes and the percentage of their occurrence appear in Fig. 1.

The results indicate that despite the fact that the participant children were of a relatively young age - which is typically related to exploratory actions - the behaviours of deliberate musicking (C5) and planning (C6) appeared slightly more than the exploratory behaviours of spontaneous musicking (C1) and sound exploration (C2).

Furthermore, a grouping of the behaviours that correspond to reflective actions (C3 and C4) and the ones that correspond to active music-making (C1, C2, C5 and C6) reveals that the "reflecting" behaviours occurred 46.18% of the total cognitive behaviours, while the active music-making behaviours occurred 53.82% (see Fig. 3). These results indicate that despite the young age of the participants, reflecting and reasoning about the musical choices appear as an integral part in children's cognitive engagement with music-making.

4.1.2. Reflection

Computer-supported music composition was selected as an open-ended task which does not include a predefined objective final "solution"; rather, it involves decision-making based on subjective criteria and self-regulated goal identification and provides the context for the emergence of a variety of processes and interactions. We identify two major findings relevant to the scope of this paper; first, despite the unstructured and the highly exploratory nature of this task, we observed that children exhibited behaviours that correspond to "making" and to "reflecting". Spontaneous and exploratory actions were mixed with deliberate actions and planning while the latter were supported by assessment and reasoning. Second, the collaborative setting of this study facilitated children's verbal interactions and negotiations during their decision-making process and consequently their reasoning and reflection on their actions. These process correspond to the mobilization of their *verbal metacognition* part of which was the *generation of explanations* during the negotiation of their task-related decisions. This means that given the opportunity (in this case collaborative setting), children as young as 5 years old actively engage in self-initiated reflection on their actions and imagine the future outcomes while being able to explain their reasoning to the collaborator. However, we observed that they often lacked the verbal abilities and the terminology for accurate explanations. For this reason, they mobilised other available modalities, such as gestures, and used the affordances of the graphical user interface of the tool provided to complement their explanatory behaviours.

4.2. Study 2: An indication for the Aha! moment

The goal of this experiment was to test the impact of the type of a robot intervention on children's problem-solving process. We used the cognitive task of the Tower of Hanoi (ToH) [31] which is used to measure children's planning abilities and inhibitory control. To reach the optimal solution, it requires participants to involve inhibition of impulsive moves that superficially bring the child closer to the goal, but are unhelpful for the longer-term solution [32]. We designed a experiment with three phases: a baseline (single child), an intervention (manipulation of the robot's behaviour) and an evaluation (child's voluntary interaction with the robot) for $N = 20$ children 5 to 7 years old. For the intervention phase, we had two conditions; in Condition1, the robot and the child solved the task in a turn-taking setting and in Condition2 we designed a child-initiated voluntary interaction with the robot. In this paper, we focus on a single child's problem-solving process to explore behavioural proper-

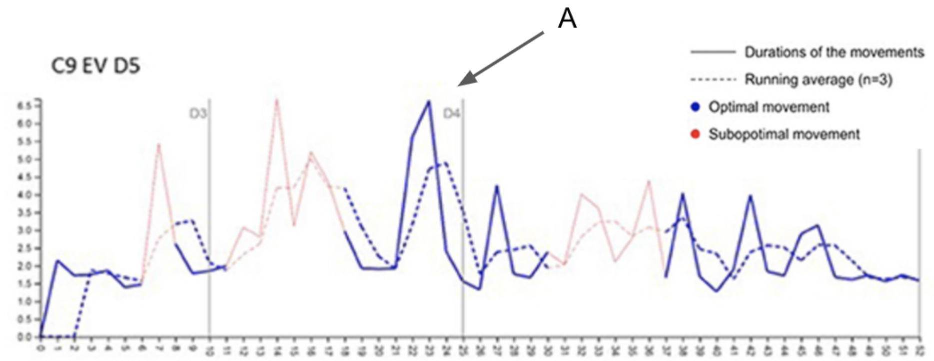

Figure 4: A child's performance of the Tower of Hanoi task over time (in seconds) (x-axis) with the duration of each move (in seconds), in addition to a moving average of the last three movements (y-axis). We observe that throughout the task the child exhibited a mixture of optimal (blue lines) / suboptimal (pink lines) and slow (lower) / fast (higher) moves.

ties relevant for our understanding of the transition from exploratory actions to the transformation of the problem representation, which requires the involvement of inhibitory control and the stabilization of the optimal performance. The details of the study appear in [33].

4.2.1. Data analysis

We evaluated the task performance in relation to the trajectory of optimal and suboptimal movements over the course of the task. The optimal movements are defined as the ones that lead to the solution of the task with the minimum number of movements. In addition, we measured the relevant speed of the movements in relation to the baseline of each participant. Given the assumption that during the task the children sustained the necessary attention, we identify point A in Fig. 4 as the point that separates the phase of mostly suboptimal moves (red peaks) with the phase of mostly optimal movements (blue peaks), which are also carried out faster.

4.2.2. Reflection

We observed exploratory behaviours that typically were characterised by increased number of suboptimal moves. We identify as an Aha! moment, the point when a transformation of the mental representation of the problem occurs which, in this task, is behaviourally manifested by the mobilization of inhibition as a strategy for the optimal solution of the task, meaning that the child inhibits the impulsive move and performs the less obvious one that will lead to the *stabilization of optimal solution of the task* (see point A in Fig. 4). This is a cognitive strategy that in the age-group of the present studies does not appear intuitively. As shown in Figure 1, behavioural properties that appear in the problem-solving process in the context of the given tasks include the performance instability, the incremental optimization and the performance stabilization. During the exploratory phase, the children were reinforced by the results of their actions which eventually guided them to the restructuring of the problem representation and consequently the use of the strategy which is based on inhibitory control. After the Aha! moment, we observe a stabilisation of the optimal moves which indicates learning. One of the limitations of this study was the fact that it was not designed in a way to facilitate the child's verbalisation of their thoughts, reasoning and reflections. For this reason, we were not able to make any inferences regarding the children's reasoning, their verbal metacognition and the generation of possible explanations during the problem-solving process.

4.3. Study 3: Social Interaction and Explanations

The purpose of study 3 was to explore the role of a social robot on children's collective problem-solving and the child-child social dynamics in a setting of two children and one robot (see Fig. 5). We built upon study 2 and we used the same task, the ToH task and the same robot. We designed a controlled 2X2 experimental study with $N = 86$ children who all participated in a baseline session (without robot), an intervention (with the manipulation of robot behaviour, in terms of its cognitive reliability and expressivity) and an evaluation session (with child-initiated form of interaction) to solve the Tower of Hanoi task with an incremental difficulty level in the different experimental configurations without any expert's intervention. For the purposes of this paper, we focus on the findings on the patterns of children's social interactions and verbal negotiations and explanations during the collective task performance. The detailed research design, analysis and findings of the study appear in [22].

Figure 5: Setting of Study 3: Two children collectively solve the Tower of Hanoi task in a turn-taking or child-initiated voluntary interaction with the robot.

4.3.1. Data analysis

We observed that the setting of the study facilitated child-child social interaction and verbal reflection, reasoning and planning appeared to be an integral part of the process which was lacking from study 2. To measure the team disparity, we define social interaction, S, as the number of task-related interactions between children.

$$S = \frac{S_1 + S_2}{L}$$

where S_n with $n = 1, 2$ refers to the number of times child n addresses their peer with a task-related verbal or non-verbal (i.e. pointing and gestures) behaviours and L refers to the number of movements needed by the team to solve the task. Our analysis showed that children had a higher S rate during the sessions with the robot, namely the Intervention ($M = 0.16, SD = 0.14$) and the Evaluation ($M = 0.13, SD = 0.092$) which differed significantly from the Baseline session ($M = 0.06, SD = 0.09$) with $p = 0.08$ and $p = 0.015$. Among the verbal manifestations we identified the utterances related to planning as one of the strategies children used to negotiate for the next movement on the ToH task. We identified the balance between children in the planning of the movements, and defined a planning disparity metric, as the absolute difference in the number of interactions initiated by each child of the team: Our analysis showed that there was a significant difference in task performance ($U = 297, p < 0.001$) between teams with a balanced planning performing better ($N = 19, M = 0.51, SD = 0.40$) compared to groups with an unbalanced planning behaviour ($n = 18, M = 1,61, SD = 0.98$). In this case, planning was used as part of the explanation formation which was observed to be one of the strategies for children's negotiations in problem-solving.

4.3.2. Reflection

Children's social verbal and non-verbal interaction during the problem-solving process in Study 3 appeared catalytic for the facilitation of their task-related planning as part of explanation generation. This was more evident in the sessions with the robot. One possible explanation for this is the fact that one of the conditions involved a robot that suggested suboptimal movements. In that case, the children engaged in child-child negotiations and explanation generation to collectively take a decision for the next move. Our observations indicate that two cognitive strategies were involved in children's explanations, *planning* as a part of an *a priori* explanation of their reasoning for a certain decision and *reflecting* as a part of an *a posteriori* explanation. We need yet to analyse the association of the strategy of planning in the context of explanatory behaviours and its relation to a preceding Aha! moment. It should be noted that additional non-verbal manifestations, such as pointing and gestures, were mobilised in the cases that a child did not have the verbal maturity to formulate the planning or the explanation.

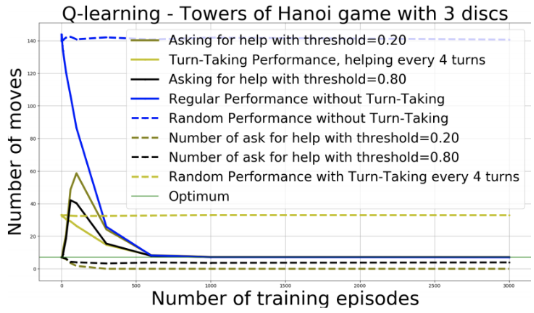

Figure 6: Asking for help scenario with different *ask for help* values: the LA2 tries to solve the game alone while being able to ask for help whenever its best action is not good enough (plot not on logarithmic scale as the agent asks for help at most 7 times)

4.4. Study 4: Multi-agent setting

This study in [34] consists of the same non-open-ended task (ToH) and collaborative setting: one learning agent (LA) and one helping agent with focus on the voluntary interaction among artificial agents. In order to explore if algorithms benefit from asking for help in collaborative problem-solving, as children do, two hypotheses are tested:

H1: Canonical interventions from an expert speed up learning.

H2: Getting help on demand from an expert accelerates finding the optimal solution compared to not on demand.

The expert intervention occurs in 2 different scenarios: 1) LA1 solves the task in collaboration with the helping agent in a “turn-taking” scenario, which results in a canonical cognitive intervention from the expert. 2)

LA2 solves the task independently, having the option to ask for help of the expert whenever (if) this is needed resulting in an *on demand* intervention. Two parameters are assessed: 1) Canonical intervention (help) rate (every 2, 3 or 4 turns), and 2) Ask-for-help threshold (from 0 to 1). The last parameter was created to simulate what happens when a child asks for help: if the best policy value is lower than the ask for help parameter, the expert will play instead of the LA.

4.4.1. Data analysis

From reinforcement learning (RL) plots, as training episodes evolve as a function of the mean number of moves to solve the task, some interpretations are extracted:

From scenario 1 it is observed that the LA is more efficient when it is helped by the expert in a turn taking scenario, and that it is even more effective when helped every 4 turns rather than every 2. The importance of exploring on its own is showcased by the agent, rather than always having the optimal solution.

Even when all approaches converge in both scenarios, in scenario 2, the agent that asks for help becomes also faster and more effective: help is most useful at the beginning of learning. After asking for help many times during the first episodes it starts solving the task by itself, resulting in an increase of inefficient moves. The agent seems to gain confidence in movements which, while imperfect, allow the task resolution by exploring different states.

Compared to the LA not being helped, the asking-for-help agent is a lot more efficient, but there is not much variation among the canonical and the help-asking configurations. This is probably due to the rather simple simulation of the trigger for the request of help. Simulating the child's behaviour is a complex task and more emphasis needs to be placed on accurately describing it, including how to simulate the "asking for help" function. Adding mechanisms such as intrinsic motivation, about the LA's desire to solve the game on its own, could make the comparison more accurate. The agent asks for help when it considers that a movement is not good enough to be played, whereas the actual mechanisms that drive the child to ask for help are more complex.

We identify emergent issues that should be part of the explanation for the design of a LA, and that are not part of the modelled problem nor a representation is accessible for the agent and thus, for its explanation:

1. The lack of a multimodal input space [35] for the agent to perform in the action-interaction loop of RL can restrain the agent from exhibiting the correct behaviour and communicating as humans would expect.
2. The lack of a natural language interactive communication interface impedes questioning the system about its confidence in real time, or uncertainty estimates.
3. The dependence of a *happy end* solving the task constraints explanations during the learning phase. How can an agent communicate to its developer or user its struggles in solving the task when it has not yet achieved a satisfactory performance? Could common AI practices (data augmentation, fine-tuning), be accessible to the agent for it to communicate and be able to choose to change them?
4. Most deep RL models rely on baselines of other agents to assess their worth, lacking comparisons in multi-agent settings [36] with children learning.
5. The inability of current RL algorithms to communicate the continual progress beyond reporting a sole reward value obtained at the end of a convergence curve makes it challenging for LAs to explain their skills to solve the credit assignment problem, their difficulties or agility to complete sub-tasks, their acting self-confidence, or learned savvy behaviours.
6. The lack of alignment of explanations of LAs with meta-learning and trustworthy AI dimensions (such as the trust calibration meta-information taxonomy [1]) should be accounted for in the explanation generation process, in the same way as mechanisms to ensure the reproducibility of insights-built explanations.

Reflection

This is the only study not involving kids, but following the same protocol as in the ToH studies. While RL model developers communicate a model learning convergence due to reaching plateaus in learning curves, these changes should as well reflect key changes in kid behaviour. However, we showed this is not always evident to map. Providing the learning agent with signals such as the Aha! moment to adjust its self learning /hyperparameter changes could be paramount to avoid blind manual engineering (on e.g., reward function crafting) processes where no common procedures exist. We believe these are the explainable dimensions that XAI for RL should work on (identifying Aha! moments, categorising problems, difficulty, environments, collaboration/competition dimensions, etc).

The difficulties to explain the learning process of a single LA could be reduced by involving interaction with other agents. One could attend to social interaction [16] and social influence as intrinsic motivation [37] learning metrics. Both showed to enhance learning in multiagent settings.

Explanations should reflect the needs for these incentives that agents depend on to progress. Once an agent learned, it is not enough that the agent performs tasks in less time, and better, but also that it uses other human factors or social outcome metrics such as in [38, 39].

5. Conclusions, Limitations and Future work

We presented a set of behavioural studies and an experiment with a Q-learning algorithm in order to identify behavioural properties that relate to the Aha! moment, i.e., the moment of the restructuring of the problem representation (Fig. 1). These behavioural properties appear as part of the transition from exploratory to explanatory behaviours (tacit to explicit knowledge). They include task-related observations such as performance instability, incremental optimization and stabilization as well as verbal metacognitive manifestations (observed only in two of our studies with children) that involve reasoning, reflection and planning. These behavioural properties seem to facilitate the Aha! moment and eventually the generation of explanations by children.

In current RL practice, detecting this moment normally translates into observing convergence curves of the objective functions being optimized (normally reaching a plateau in cumulative reward or optimized loss, usually both). This is an external signal not usually leveraged by the agent. Although there are exceptions such as the use of artificial curiosity signals for self learning of the agent [40]), in regular AI model development practice, we must highlight the need for easier mechanisms to convey actions that demonstrate the difficulties of the agent until convergence plateaus and/or a sufficient level of an XAI metric are reached.

We summarise the main points we propose to consider in approaches evaluating XAI, as follows:

1. The Aha! moment (or problem representation restructuring) acts as the intermediate step between non-explainable and explainable behaviours. In a deeper view, since the explanation acts as an interface between the model and a given target audience, the Aha! moment is a trigger signal for a model to start elaborating explanations. More effort should be put into specifying the meaning of the Aha! moment in various tasks that RL models are currently tackling. Defining and detecting high level policies characterising an Aha! moment (e.g. in terms of key/exploratory action sequences) can be signs we should be able to not only programmatically detect, but also communicate. In this way, we can achieve explainable and reliable models, since Aha! signs must act as an additional proxy to attain trustworthy systems.
2. Children understand but sometimes lack the cognitive and metacognitive skills to explain. Explanations are subject to both biological and artificial systems' understanding of properties of a given task, and in young children explanations are subject to their verbal abilities. Children often use gestures such as pointing, which means that explanations that can support human-AI interaction are subject to tools responsible for social interaction. Aiming towards human-level AI requires a broader set of key social skills for complex embodied communication in multimodal settings within constantly evolving social worlds [41].
3. The hybrid use of strategies of "planning" and "insight": Questioning the false dilemma of logical reasoning vs machine learning, we argue for a synergy between these two paradigms in order to obtain hybrid AI systems.
4. Both social interaction [16] and social influence as intrinsic motivation [37] show to be enhancers for learning in multi-agent settings.

This paper is our first attempt to synthesise the results of our research on children's problem solving in different settings and combine them with our research on XAI. However, due to space limitations, we are limited to provide overviews without in-depth analysis. We aim to tackle the latter in our future work.

We hope this work is useful beyond developmental robotics and AI, i.e., facilitating an effective and ethical deployment of RL systems, e.g. from energy building management to AI for health, where evaluating single reward functions simply does not reflect nor assess the complexity of the system nor the difficulties it has to deal with.

Future work should aim to involve more tangible evaluation metrics that both 1) optimize technical robustness more broadly, and 2) reflect a human-centered view where machine learning factors are questioned, monitored and explained in parallel ways to how children learn. Evaluation mappings across human and machine learning will allow us to better assess the trade-offs between AI assisted decision making and policies.

6. Acknowledgments

Charisi is supported by the HUMAINT project of the JRC; Díaz-Rodríguez by IJC2019-039152-I funded by MCIN/AEI /10.13039/501100011033 by "ESF Investing in your future" and Google Research Scholar Program; and Merino by Programa Operativo FEDER Andalucia 2014-2020 through the project DeepBot (PY20_00817).

References

[1] G. J. Cancro, S. Pan, J. Foulds, Tell me something that will help me trust you: A survey of trust calibration in human-agent interaction, arXiv preprint arXiv:2205.02987 (2022).

[2] A. B. Arrieta, N. Díaz-Rodríguez, J. Del Ser, A. Bennetot, S. Tabik, A. Barbado, S. García, S. Gil-López, D. Molina, R. Benjamins, et al., Explainable artificial intelligence (xai): Concepts, taxonomies, opportunities and challenges toward responsible AI, Information fusion 58 (2020) 82–115.

[3] J. Kounios, M. Beeman, The aha! moment: The cognitive neuroscience of insight, Current directions in psychological science 18 (2009) 210–216.

[4] H. Stuyck, A. Cleeremans, E. Van den Bussche, Aha! under pressure: The aha! experience is not constrained by cognitive load, Cognition 219 (2022) 104946.

[5] L. B. Smith, L. K. Slone, A developmental approach to machine learning?, Frontiers in psychology 8 (2017) 2124.

[6] B. M. Lake, T. D. Ullman, J. B. Tenenbaum, S. J. Gershman, Building machines that learn and think like people, Behavioral and brain sciences 40 (2017).

[7] P.-Y. Oudeyer, L. B. Smith, How evolution may work through curiosity-driven developmental process, Topics in Cognitive Science 8 (2016) 492–502.

[8] M. M. Andersen, J. Kiverstein, M. Miller, A. Roepstorff, Play in predictive minds: A cognitive theory of play (2021).

[9] J. S. Bruner, Organization of early skilled action, Child development (1973) 1–11.

[10] P. Barrouillet, Theories of cognitive development: From piaget to today, Developmental Review 38 (2015) 1–12.

[11] M. H. Siegel, R. W. Magid, M. Pelz, J. B. Tenenbaum, L. E. Schulz, Children's exploratory play tracks the discriminability of hypotheses, Nature communications 12 (2021) 1–9.

[12] J. Chu, L. E. Schulz, Play, curiosity, and cognition, Annual Review of Developmental Psychology 2 (2020) 317–343.

[13] A. Gopnik, Childhood as a solution to explore–exploit tensions, Philosophical Transactions of the Royal Society B 375 (2020) 20190502.

[14] M. Pelz, C. Kidd, The elaboration of exploratory play, Philosophical Transactions of the Royal Society B 375 (2020) 20190503.

[15] J. M. Unterrainer, A. M. Owen, Planning and problem solving: from neuropsychology to functional neuroimaging, Journal of Physiology-Paris 99 (2006) 308–317.

[16] K. Ndousse, D. Eck, S. Levine, N. Jaques, Learning social learning, Internation Conference on Machine Learning (ICML) (2021).

[17] L. S. Vygotsky, M. Cole, Mind in society: Development of higher psychological processes, Harvard university press, 1978.

[18] C. Colas, T. Karch, C. Moulin-Frier, P.-Y. Oudeyer, Vygotskian autotelic artificial intelligence: Language and culture internalization for human-like AI, arXiv preprint arXiv:2206.01134 (2022).

[19] J. Lindblom, T. Ziemke, Social situatedness of natural and artificial intelligence: Vygotsky and beyond, Adaptive Behavior 11 (2003) 79–96.

[20] D. Wood, J. S. Bruner, G. Ross, The role of tutoring in problem solving., Child Psychology & Psychiatry & Allied Disciplines (1976).

[21] L. Kerawalla, D. Pearce, J. O'Connor, R. Luckin, N. Yuill, A. Harris, Setting the stage for collaborative interactions: Exploration of separate control of shared space., in: AIED, 2005, pp. 842–844.

[22] V. Charisi, L. Merino, M. Escobar, F. Caballero, R. Gomez, E. Gómez, The effects of robot cognitive reliability and social positioning on child-robot team dynamics, in: 2021 IEEE International Conference on Robotics and Automation (ICRA), IEEE, 2021, pp. 9439–9445.

[23] T. I. Lubart, C. Mouchiroud, Creativity: A source of difficulty in problem solving, The psychology of problem solving (2003) 127–148.

[24] W. Carpenter, The aha! moment: The science behind creative insights, in: Toward Super-Creativity-Improving Creativity in Humans, Machines, and Human-Machine Collaborations, IntechOpen, 2019.

[25] Y. Chu, J. N. MacGregor, Human performance on insight problem solving: A review, The Journal of Problem Solving 3 (2011) 6.

[26] T. Miller, Explanation in artificial intelligence: Insights from the social sciences, Artificial intelligence 267 (2019) 1–38.

[27] M. Shvo, T. Q. Klassen, S. A. McIlraith, Towards the role of theory of mind in explanation, in: International Workshop on Explainable, Transparent Autonomous Agents and Multi-Agent Systems, Springer, 2020, pp. 75–93.

[28] R. K. Yin, et al., Design and methods, Case study research 3 (2003).

[29] V. Charisi, C. C. Liem, E. Gomez, Novelty-based cognitive processes in unstructured music-making settings in early childhood, in: 2018 Joint IEEE 8th International Conference on Development and Learning and Epigenetic Robotics (ICDL-EpiRob), IEEE, 2018, pp. 218–223.

[30] R. S. Siegler, icrogenetic analyses of learning, in: Handbook of child psychology: Cognition, perception, and language, John Wiley & Sons Inc, 2006, p. 464–510.

[31] H. A. Simon, The functional equivalence of prob-

lem solving skills, Cognitive psychology 7 (1975) 268–288.
[32] N. A. Zook, D. B. Davalos, E. L. DeLosh, H. P. Davis, Working memory, inhibition, and fluid intelligence as predictors of performance on tower of hanoi and london tasks, Brain and cognition 56 (2004) 286–292.
[33] V. Charisi, E. Gomez, G. Mier, L. Merino, R. Gomez, Child-robot collaborative problem-solving and the importance of child's voluntary interaction: a developmental perspective, Frontiers in Robotics and AI 7 (2020) 15.
[34] A. Bennetot, V. Charisi, N. Díaz-Rodríguez, Should artificial agents ask for help in human-robot collaborative problem-solving?, Brain-PIL Workshop at ICRA, Paris/Remote (2020).
[35] A. Holzinger, M. Dehmer, F. Emmert-Streib, R. Cucchiara, I. Augenstein, J. Del Ser, W. Samek, I. Jurisica, N. Díaz-Rodríguez, Information fusion as an integrative cross-cutting enabler to achieve robust, explainable, and trustworthy medical artificial intelligence, Information Fusion 79 (2022) 263–278.
[36] A. Heuillet, F. Couthouis, N. Díaz-Rodríguez, Explainability in deep reinforcement learning, Knowledge-Based Systems 214 (2021) 106685.
[37] N. Jaques, A. Lazaridou, E. Hughes, C. Gulcehre, P. Ortega, D. Strouse, J. Z. Leibo, N. De Freitas, Social influence as intrinsic motivation for multi-agent deep reinforcement learning, in: International Conference on Machine Learning, PMLR, 2019, pp. 3040–3049.
[38] J. Perolat, J. Z. Leibo, V. Zambaldi, C. Beattie, K. Tuyls, T. Graepel, A multi-agent reinforcement learning model of common-pool resource appropriation, Advances in Neural Information Processing Systems 30 (2017).
[39] A. Heuillet, F. Couthouis, N. Díaz-Rodríguez, Collective eXplainable AI: Explaining Cooperative Strategies and Agent Contribution in Multiagent Reinforcement Learning With Shapley Values, IEEE Computational Intelligence Magazine 17 (2022) 59–71.
[40] D. Pathak, P. Agrawal, A. A. Efros, T. Darrell, Curiosity-driven exploration by self-supervised prediction, in: International conference on machine learning, PMLR, 2017, pp. 2778–2787.
[41] G. Kovač, R. Portelas, K. Hofmann, P.-Y. Oudeyer, SocialAI: Benchmarking socio-cognitive abilities in deep reinforcement learning agents, arXiv preprint arXiv:2107.00956 (2021).