Implementing Artificial Intelligence across Task Types: Constraints of Automation and Affordances of Augmentation

Structured abstract

Purpose: Uncovering the constraints of automation and the affordances of augmentation related to implementing AI-powered systems across different task types: mechanical, thinking, and feeling.
Design/methodology/approach: Qualitative study involving 45 interviews with various stakeholders in artistic gymnastics, for which AI-powered systems for the judging process are currently developed and tested. Stakeholders include judges, gymnasts, coaches, and a technology vendor.
Findings: We identify perceived constraints of automation, such as too much mechanization, preciseness, and inability of the system to evaluate artistry or to provide human interaction. Moreover, we find that the complexity and impreciseness of the rules prevent automation. In addition, we identify affordances of augmentation such as speedier, fault-less, more accurate and objective evaluation. Moreover, augmentation affords providing an explanation, which in turn may decrease the number of decision disputes.
Originality/value: Our granular approach provides a novel point of view on AI implementation, as our findings challenge the notion of full automation of mechanical and partial automation of thinking tasks. Therefore, we put forward augmentation as the most viable AI implementation approach. In addition, we developed a rich understanding of the perception of various stakeholders with a similar institutional background, which responds to recent calls in socio-technical research.
Research limitations: While the unique context of our study is revealing, the generalizability of our specific findings still needs to be established. However, the approach of considering task types is readily applicable in other contexts.
Practical implications: Our research provides useful insights for organizations that consider implementing AI for evaluation in terms of possible constraints, risks, and implications of automation for the organizational practices and human agents while suggesting augmented AI-human work as a more beneficial approach in the long term.

Keywords: Artificial intelligence; Affordances and constraints; Automation; Augmentation; Task types; Evaluation; Sports digitalization

Implementing Artificial Intelligence across Task Types: Constraints of Automation and Affordances of Augmentation

1. Introduction

Artificial Intelligence (AI) is considered to be a general purpose technology that can be adopted in various work processes, across industries and functions (e.g., marketing, human resources, operations) (Benbya *et al.*, 2021; Rai *et al.*, 2019). Implementing AI-powered systems in work processes may have far-reaching implications, for instance in terms of changing roles of human specialists, new forms of cooperation and decision-making, re-education of workers, and labor force reductions (Huang and Rust, 2018; Yu *et al.*, 2023). Notably, there are important differences in the ways in which AI-powered systems can be implemented in work processes, typically designated as automation versus augmentation. Automation refers to AI-powered systems completely taking over tasks, keeping humans out of the process (Dellermann, Ebel, *et al.*, 2019). Potential advantages stem from performance increases, with capabilities that (far) exceed those of humans (Seidel *et al.*, 2019), as well as cost and time gains (Ebel *et al.*, 2021). Instead, augmentation refers to AI serving as a tool supporting humans' sensorial and cognitive capabilities with potential for mutual learning, through the exchange of skills and knowledge between humans and AI systems (van den Broek *et al.*, 2021; Raisch and Krakowski, 2021).

In our reading of the literature, we perceive a tendency to highlight various limitations of AI automation, for instance related to workers' productivity, skills, tasks, overall psychological well-being, and their confidence in a future career (Luo *et al.*, 2019a; Tong *et al.*, 2021), which then leads to recommendations for AI augmentation, accompanied by changes in the work processes and human tasks (Daugherty and Wilson, 2018; Dellermann, Ebel, *et al.*, 2019). Against this background, we argue that it is important to develop a more nuanced understanding and we consider the level of task types to contribute to the discourse (Huang and Rust, 2024; Huang *et al.*, 2019). In particular, based on three task types (mechanical, thinking, feeling), we provide granular insight into the constraints of automation and the affordances of augmentation. Such insight is critical, because each approach involves distinct and sometimes even conflicting measures for implementation (Benbya *et al.*, 2021).

Moreover, we draw from a technology affordance lens to provide “a new way of seeing things” (Volkoff and Strong, 2017, p. 9) in a socio-technical relationship (Yu *et al.*, 2023). In keeping with previous research, we examine affordances both in their enabling and constraining capacity (Autio *et al.*, 2018). Technology affordances refer to the possibilities of action that individuals or organisations, within their context and with particular purposes, perceive in a technology (Leonardi, 2023; Strong *et al.*, 2014). The affordance lens is particularly relevant to study AI, as it not only helps to “theorize how [AI] agency affects us, as humans, in the context of our work,” but more generally to consider “how agency materializes at the human-machine-institution interface” (Leonardi, 2023, p. xvii). Indeed, the affordance lens is especially useful when: the focus is on the relationship between actors, technology, and institutions (Leonardi, 2023; Strong *et al.*, 2014); multiple actors with a variety of potential goals and actions are involved (Autio *et al.*, 2018); and perceptions are examined (Felin *et al.*, 2016). Moreover, we do not limit perceptions to how technology is currently used, but also include imagined affordances and constraints that capture attitudes and expectations anticipated for the future (Brooks and Saveri, 2017; Nagy and Neff, 2015). Hence, the research question we aim to address in this paper is: *“What are constrains of automation and affordances of augmentation of AI implementation across task types?”*

To address this question, we searched for an information-rich, empirical setting that is undergoing a transition from pure human-based work processes to one in which AI is implemented. We found this setting in artistic gymnastics, in which currently an AI-powered system for the judging process is developed, yet it is unclear whether it would be a system replacing (automation) or supporting (augmentation) human judges (Fujiwara and Ito, 2018). We conducted an exploratory case study and interviewed 45 stakeholders from various countries who are professionally active in this context, with the goal of obtaining a holistic understanding of their perceptions on AI implementation for the work processes involved in artistic gymnastics judging. To do so, we categorized the different tasks of judges in terms of their mechanical, thinking, and feeling characteristics (Huang *et al.*, 2019) and identified constraints of automation and affordances of augmentation on that basis.

This paper adds to socio-technical research on AI implementation in several ways. First, by considering constraints of automation and affordances of augmentation across task types, we obtain granular insight into the drivers and inhibitors of these two approaches to AI implementation. (Baer *et al.*, 2022; Raisch and Krakowski, 2021). Second, our approach considers potentially contradictory viewpoints of various stakeholders (Sarker *et al.*, 2019), which can result in the successful tailoring and integration of the technological changes to the social and institutional context (Leonardi, 2023). By doing so, the technology adoption and use can increase and the wellbeing and job satisfaction of the workers involved can improve (Sarker *et al.*, 2019; Yu *et al.*, 2023). Third, our empirical setting is real-life, which is different from prior studies conducted in "laboratory settings" that "disregard the role of humans, and the wider [...] societal implications" (Raisch and Krakowski, 2021, p. 203).

While the sports context can be considered unique because of the specific stakeholders involved (Stremersch *et al.*, 2022), this context has extensively been employed to study phenomena in management research in general (Day *et al.*, 2012; Marino *et al.*, 2015) and in information systems in particular (e.g., Jarvenpaa and Standaert, 2018; Xiao *et al.*, 2017). This can be related to the extreme focus on high performance (management) (Fonti *et al.*, 2023) and the specificities of digitalization when there is inherent physicality (Goebeler *et al.*, 2021). For practitioners, our work points to the significance of considering task types as well as the perceptions of various types of stakeholders when deciding how to implement an AI-powered system. Indeed, in their pursuit of greater productivity, efficiency, and profitability, organizations often introduce AI-powered systems into work processes without full consideration of the various implications and risks involved (Agrawal *et al.*, 2017).

We proceed as follows. In section two, we review the literature on the two main approaches to implementing AI-powered systems and on the role of task types. In the third section, we present our research context and framework and the fourth section outlines our methodology. In section four, our qualitative findings are presented across the two approaches (automation or augmentation) and three task types. We discuss the implications of our findings for automation and augmentation of work in section five and conclude the paper in section six.

2. Literature Review

The introduction of AI-based systems in organizations' work processes raises the question of which implementation approach is suitable. In this section, we discuss the literature on two key approaches, namely automation and augmentation and we close the section with a review of AI implementation across different types of tasks.

2.1 Automation

Researchers have drawn attention to many issues related to the automation of work processes, in which AI systems substitute for humans (Seidel *et al.*, 2019). A key advantage of automation is that AI taking over tasks of human workers can result in more effective information-processing and decision-making. In addition to enhancing performance, advantages of automation include overcoming human cognitive limitations and reducing costs (Benbya *et al.*, 2021). For instance, Luo et al (2019) found that AI chatbots were as effective as experienced workers in (highly structured) outbound sales calls. However, organizations may be confronted with a dilemma, because disclosing the use of AI, in interactions with customers or employees, may lead to negative perceptions and lower productivity (Luo *et al.*, 2019b; Tong *et al.*, 2021). Recent work has provided more nuance, by stating that there are no 'human-free' processes and that "automation should not mean that humans are out-of-the-loop" completely (Constantinides *et al.*, 2024, p. 16). More specifically, the authors argue that, in tasks that involve high uncertainty, a dynamic model emerges that based on different types of human and AI learning involves interactions between uncontrolled and controlled automation and between limited and expanded automation (Constantinides *et al.*, 2024).

Despite potential benefits of automation, AI fundamentally relies on rationality based on impersonal, quantitative calculations of big sets of data. Indeed, AI lacks a capacity to share and acquire knowledge via engaging with the organizational environment and learning from the experience that is inherent in human learning (Balasubramanian *et al.*, 2022). Completely substituting human decision-making with AI decreases the richness of the organizational learning and background knowledge, hence it ignores environmental changes and limits "within-organizational diversity in routines" (Balasubramanian *et al.*, 2022, p. 448). Also, this can exacerbate learning myopia associated with the limitations of human learning, namely through focusing on the short-term problems with easier solutions and thus ignoring the long-term perspectives. In turn, the machine learning myopia is explained by the inclination to learn from well-defined historical data and ignore variants that do not fit with them, as well as by human myopia previously built into the available data (Balasubramanian *et al.*, 2022). In summary, pursuing automation of a process may provide productivity benefits in the short term, but by doing so preclude an organization from making more drastic changes that could lead to even more benefits in the long term (Daugherty and Wilson, 2018).

Furthermore, scholars have not reached consensus on the long-term implications of automation. Some believe automation will benefit society overall, increase productivity, propel economic growth, and create more jobs, in particular in the emerging field of maintaining, evaluating, and explaining AI-powered systems (Seidel *et al.*, 2019; Wilson *et al.*, 2017). Others predict that many workers will face harm from the automation of labor in the coming decades (Agrawal *et al.*, 2017), leading to extensive job losses and social inequality rising further due to unemployment (Brynjolfsson *et al.*, 2017). Experts in this camp argue that, while automation may lead to new jobs, these will be either undesirable, low-paying ones or system-support-related jobs that require advanced qualifications that cannot be performed by those who lose their job due to automation (Wilson *et al.*, 2017).

For individuals, automation may lead to (partial) unemployment, decreased work quality, and problems related to social and financial instability (Davenport and Ronanki, 2018). Together, these consequences of automation could highly destabilize workers' situations. They might lose the required professional skills, be uncertain as to their professional perspective, and find themselves unable to plan their career or life in general (Crawford and Whittaker, 2016). It seems that, while automation in work processes can lead to benefits mostly for organizations, the consequences that ensue for the individual human workers may be negative and irreversible (Yu *et al.*, 2023). Given such "automation anxiety" (Baer *et al.*, 2022), would augmentation be a better approach?

2.2 Augmentation

Augmentation has been studied by different fields, but there is no commonly accepted definition yet (Baer *et al.*, 2022). Augmentation is often referred to in contrast to automation (Raisch and Krakowski, 2021), as it involves both human and AI agents that interact to perform tasks (Rai *et al.*, 2019). Such interaction or collaboration often lies at the center of defining augmentation, but another approach is to consider “what is augmented with AI?” (Baer *et al.*, 2022). Some of the most common views are that AI can augment performance, decision-making and problem solving, and cognitive capacity and skills (Baer *et al.*, 2022).

Furthermore, augmentation is characterized by continuous mutual learning and intense knowledge exchange between humans and machines that enables improvements to both human and machine capabilities and skills (van den Broek *et al.*, 2021; Constantinides *et al.*, 2024). In this case, AI can be used for ‘real-time’ processing and generating information before passing it on to human workers, resulting in potentially large time savings for the overall work processes (Dellermann, Ebel, *et al.*, 2019). The role of humans can therefore be to define and set goals, and exercise overall decision-making control. For example, designers and engineers can set parameters for AI tools to augment the search for design alternatives that the algorithms can quickly generate, and then make final decisions based on that input (Seidel *et al.*, 2019). Likewise, AI can take over the routine initial phase of sales lead generation, after which humans can take over and focus on the more value-adding, creative part of the job (Jia *et al.*, 2024). Additionally, human experts can monitor the application of moral values and ethical compliance or explain decisions made by AI agents to other stakeholders that do not have technical knowledge about the “black-box” of AI (Asatiani *et al.*, 2021). As such, humans can help increase the transparency of AI work processes and take accountability for the end result (Berente *et al.*, 2021; Wilson *et al.*, 2017).

AI augmentation requires organizations to rethink their business processes and rethinking roles, which may require humans "to do new and different things” and "to do things differently" (Wilson and Daugherty, 2018, p.123). In particular, employees will need to develop "fusion skills" to learn to work effectively with machines and delegate tasks to them, explain the outcomes of the AI decision-making process, train intelligent agents, and sustain responsible, legal, and ethical work of AI (Wilson and Daugherty, 2018). Augmentation also allows more qualified and experienced workers to focus on more complex problem-solving, freeing up their cognitive resources for finding more useful, efficient and creative solutions at work, which in turn may positively impact their psychological well-being (Wang *et al.*, 2024).

However, the concerns about the well-being and welfare of low-skilled workers remain. Indeed, low-skilled workers tend to show lower productivity and benefit less from cooperation with AI due to their inability to take advantage of the opportunities presented by AI collaboration and their unwillingness to create new ways and scenarios for their work (Wang *et al.*, 2024). Moreover, employees with more seniority demonstrate a higher degree of sensitivity to AI's mistakes, a lower degree of trust in it, and a higher degree of resistance to it. In sum, less qualified and more senior workers, demonstrate greater stress, more tension, lower morale, and a stronger fear of being replaced when collaborating with AI (Balasubramanian *et al.*, 2022; Jia *et al.*, 2024).

Notwithstanding the potential benefits of an augmentation-oriented approach, some disadvantages may prevent successful implementation. Firstly, while augmentation affords to overcome a machine's limitations via humans' intuition and other attributes, human biases and subjectivity can sometimes carry over to machines and influence their decision-making (Benbya *et al.*, 2021). Secondly, building solid co-operation between humans and machines requires specific measures and (financial) resources, for establishing coordination and knowledge exchange between human and machine agents (Raisch and Krakowski, 2021).

In summary, this approach enables exploiting the strengths of both humans and AI and compensating their shortcomings for affording faster, more efficient, and highly accurate evaluation and decision-making and reaching superior productivity (Dellermann, Ebel, *et al.*, 2019). However, the appropriate extent of human involvement may depend on the task characteristics (Constantinides *et al.*, 2024; Seidel *et al.*, 2019), which is what we discuss next.

2.3 Task Types

Scholars and practitioners agree that when discussing AI implementation, the level of jobs is too broad, but rather tasks should be considered (Davenport and Ronanki, 2018). Dimensions of the nature of tasks that have been considered include the extent to which a task is routine, abstract, complex, analytical, and involves structured data (Baer *et al.*, 2022; Dellermann, Calma, *et al.*, 2019; Rinta-Kahila *et al.*, 2023). Recent work has studied tasks with a low error margin (e.g., medical decision making) and with a varying extent to which information can be used to reduce uncertainty (Constantinides *et al.*, 2024).

However, a recent framework of task types that has received significant traction in academia since it was proposed (Huang and Rust, 2018) and that has also been updated (Huang and Rust, 2024; Huang *et al.*, 2019) distinguishes three task types: mechanical, thinking, and feeling. Mechanical tasks are characterized by their standardized, repetitive, routine, and simple nature. Thinking tasks involve systematic, rule-based, and well-defined but potentially complex processing, evaluation, and analysis of data and information, which is the reason why they often require experts to be involved. Tasks involving social communication and interaction and tasks that involve empathy and emotions, belong to feeling tasks. While most tasks involve some combination of the three types, usually one type can be identified as the primary one. Moreover, a work process usually consists of various combinations of the type of tasks described here and can be split accordingly (Huang *et al.*, 2019).

The development of AI capabilities involves different stages and progress has evolved from mechanical capabilities in the past, over thinking at the moment, to feeling in the future (Huang *et al.*, 2019). When transferring to a new level of intelligence, an AI system retains the capabilities of the previous level, for instance, thinking intelligence also encompasses mechanical capabilities (Huang and Rust, 2024). The mechanical intelligence of AI is well studied and widely used in, for example, factory automation. In the process of implementation of mechanical tasks, the AI's capabilities to learn and adapt are limited. To implement thinking tasks, such capabilities are much more important, and eventually can evolve to a level exceeding human cognitive capabilities.

Finally, feeling intelligence was long believed to be the exclusive prerogative of humans, but recent AI developments have focused on this domain as well (Song *et al.*, 2022). Moreover, generative

AI potentially represents an inflection point in this area (Huang and Rust, 2024). Indeed, generative AI intends to mimic human interaction and communication and includes such capabilities as emotion recognition, generation of empathic responses and providing recommendations for resolving users' potential emotional challenges (Huang and Rust, 2024). Therefore, generative AI is considered a tool organizations can use to for instance build strong and long-term relationships with customers, based on establishing personalized and emotion-laden communication. Nevertheless, current (generative) AI systems can still lack accuracy in emotion recognition, communication, and management due to ambiguity in input data or due to limited commonsense (Wang *et al.*, 2024). Therefore, widespread application of AI for feeling tasks may still be a few years, if not decades, away. Certainly, for the foreseeable time, feeling AI still requires oversight and verification by human experts for it to further develop (Huang and Rust, 2024).

3. Research Context

In this section, we describe the context in which we conducted our empirical study, which is specific and unique (Stremersch *et al.*, 2022). Our choice for the field of artistic gymnastics was due to the ongoing developments of an AI-powered judging system for the sport.

3.1 Gymnastics Judging as Mechanical, Thinking, and Feeling Tasks

First, we briefly introduce artistic gymnastics. This sport involves different apparatus, with some differences for women's and men's competitions. Competitions consist of gymnasts each in turn performing a routine, which consists of several elements. The gymnasts are evaluated based on their performance, following the Code of Points (CoP). Indeed, for each apparatus, the CoP consists of a list and description of all obligatory and secondary elements included in a routine, all technical details of how each element should be executed (e.g., the length, height, and angles of the jumps, rotations, and movements), and the number of points assigned for each element, as well as the deductions for errors made.

During international competitions, judges need to record all scores and deductions in symbol notation on evaluation sheets, all the while watching the routine of an athlete that can consists of a rapid sequence of elements. This requires a very high concentration and a very good "programmatic" knowledge of the symbol notations for the different elements, of which there are hundreds. Moreover, there are many similar-looking symbols and not all judges use them in the same way. Subsequently, the scores of the six judge (per apparatus) have to be inputted into a computer. Then, the two extreme (minimum and maximum) scores are eliminated and the average of the four remaining scores is computed and presented to the athletes and fans (on the big screen). These scores can be challenged, for which the gymnast's coach has to submit an inquiry to the judges asking to revise the score and the federation has to pay for this.

In addition, judges contribute to the athletes' training at different stages: during national training camps and during the last training prior to the actual competitions, which is called "podium training." During the training camps, judges of a country provide general advice to the athletes and coaches. During podium training, judges might give advice on how to improve the athletes' routine or what potential errors they see beforehand so that the athletes can change something in the routine at the last minute before the competition and increase their chances for higher scores.

We have analyzed the judging work conducted by the (human) panel of judges according to the task types presented above: mechanical, thinking, and feeling. We did this by examining the list of tasks and responsibilities presented in the international federation's judges' rulebook (2022-2024)[1] and aimed to be both complete (including judges' tasks outside of the actual competition) and granular (splitting up tasks into singular elements). We then categorized these task elements according to the primary type of capabilities involved, considering that tasks are usually not purely mechanical, thinking, or feeling, but rather a combination of some or all of them (Huang *et al.*, 2019). The resulting categorization can be found in Table 1, while the intermediate step with detailed explanations is provided in Appendix 1.

[1] Available via: https://www.gymnastics.sport/site/rules/

Table 1. Categorization of judging tasks

Mechanical tasks	Thinking tasks	Feeling tasks
Related to handling the scores: • Monitoring exact measures of the routine (e.g., duration, height) • Recording scores in symbol notation on evaluation sheets • Submitting evaluation sheets • Entering the individual scores into a computer **Managing the technical infrastructure (technical committee):** • Verifying the parameters of the apparatus • Adjusting the parameters of the apparatus	**Related to judging:** • Internalizing the rules • Applying the rules • Monitoring and coordinating the application of the rules (for superior judges) • Helping the athletes understand the rules (training) • Updating the rules	**Related to judging:** • Being objective and fair • Evaluating artistry and creativity **Related to communication:** • Resolving possible conflicts during the competitions • Motivating, supporting, and guiding the gymnasts before actual competitions • Non-verbal communication with the athletes • Communication with the audience and media

(Source: Authors' own creation)

3.2 An AI-Powered System for Judging in Artistic Gymnastics

In 2016, a joint project for the development and testing of an AI-powered judging system was initiated by three parties: technology provider Fujitsu, the International Gymnastics Federation and the International Olympic Committee. Due to its advanced technical capabilities, the AI-based judging system is expected to resolve or at least mitigate some limitations of the existing human-based judging system (e.g., cognitive limitations and human bias) and provide a more accurate, transparent, objective, and faster process for judging the gymnasts' routines. In case of successful development and implementation of this system, the goal is to use it for other judged sports as well (e.g., figure ice skating). Artistic gymnastics was chosen by the developers of the system as the starting field because of the complexity of the judging rules (cf. CoP). The system under development combines different advanced technologies, such as 3D laser sensors and AI-based joint-position-recognition (Fujiwara and Ito, 2018), to capture the gymnasts' movements and skeletal motion in 360 degrees. Then, the system compares the obtained data with the visual representation of the elements built into the database of the system based on the CoP, in order to indicate mistakes and provide the number of points that should be deducted to obtain the final score (Fujiwara and Ito, 2018).[2]

[2] More information about the system can be found at https://medium.com/syncedreview/meet-fujitsus-ai-gymnastics-judges-8cb52613b2a

4. Methodology

In this section, we outline our qualitative research approach and our methodology for collecting and analyzing the interviews.

4.1 Qualitative Data Collection

In keeping with the exploratory nature of our research, we conducted an in-depth case study (Yin, 2014) in order to "rely more on induction than deduction in understanding a phenomenon" (Sarker *et al.*, 2018, p. 914). In an exploratory study, expert interviews can be an effective way to generate insight, as experts "possess specific knowledge that relates to a clearly demarcated range of problems" and they "exert influence by establishing a particular issue-framing" (Bogner *et al.*, 2018, pp. 655–656). To select the expert informants, we employed purposeful sampling (Patton, 2002) with the inclusion criterion that they have an international level of qualification in artistic gymnastics (i.e., participating in international competitions) and would be directly affected by the AI-powered system's implementation. We conducted 45 semi-structured interviews, namely with 20 judges, 11 gymnasts, 8 coaches of international teams, 2 technical directors in artistic gymnastics, 2 representatives of the International Federation of Gymnastics, and 2 representatives of Fujitsu (the company developing the system). An overview of the informants is presented in Table 2.

Table 2. The informants

Judges		Gymnasts		Coaches		Others		
No.	Pseudonym	No.	Pseudonym	No.	Pseudonym	No.	Pseudonym	Role
1	Abby	1	James	1	Michael	1	Steven	Tech. director
2	Bella	2	John	2	William	2	Mary	Tech. director
3	Charlie	3	David	3	Daniel	3	Simon	FIG
4	Edward	4	Thomas	4	Kyle	4	Adam	FIG
5	Felicity	5	Mark	5	Margaret	5	Joona	Fujitsu
6	Harry	6	Lauren	6	Jessica	6	Caleb	Fujitsu
7	Lilly	7	Reece	7	Paul			
8	Nick	8	Fabian	8	Kevin			
9	Norman	9	Nathan					
10	Sarah	10	Jacob					
11	Ulla	11	Damian					
12	Tracy							
13	Isabella							
14	Mia							
15	Sofia							
16	Emily							

17 Bob 18 Don 19 Katarina 20 Josh			

(Source: Authors' own creation)

In the interviews, we aimed to collect a broad base of expert opinions and used open questions designed to encourage the participants to share their opinions on the various topics related to the research question (see example questions in Appendix 2). As the development and testing of the system was ongoing during the data collection, we examine perceptions and expectations with regards to the potential uses and changes that the technology introduces in terms of achieving or constraining desired outcomes (Essén and Värlander, 2019). This has been referred to as projective agency, which “encompasses the imaginative generation of actors of possible future trajectories of action… in relation to actors’ hopes, fears, and desires for the future” (Emirbayer and Mische, 1998, p. 971). A related notion is that of imagined affordances, which “emerge between users’ perceptions, attitudes, and expectations” and are deemed especially relevant for studying algorithms (Nagy and Neff, 2015, p. 1). All interviews were tape-recorded and transcribed. All non-English-language interviews were translated into English. In keeping with our promise to present the data anonymously, we use pseudonyms in this paper.

4.2 Data Analysis

We conducted the analysis of all interview material by using software designed for qualitative data analysis (ATLAS.ti). The interview data was analyzed via three coding techniques: 1) open coding, 2) axial coding, and 3) selective coding. Our detailed coding scheme is provided in Appendix 3.

First, open coding enable us to conduct initial analyses of the data and identify ideas, patterns, and opinions among the participants. This step was both inductive and deductive, in particular with regard to designating the codes as relating to a human panel of judges or an AI-powered system and with regard to the two AI implementation approaches (automation and augmentation). As to the latter, automation was referred to as "replacement" or "substitution" in the interviewees' quotes, and augmentation as "support", "supporting system" or "help". At this stage, we had 46 distinct codes, examples of which are provided in Appendix 4.

Second, we used axial coding to identify the relationships between common opinions and perceptions related to constraints of automation and affordances of augmentation of an AI-powered system. Identification of similarities and patterns was performed across interviewee groups, yet we did not observe the stakeholder group to be a distinguishing factor in opinions. In this stage, 12 constraints of automation were defined and 10 affordances of augmentation.

Third, to address our research question, selective coding was used to deduct the core themes. While these themes are grounded in the data (through the previous coding steps), this step also involved a deductive angle as the pre-defined task types were used to categorize themes, while also maintaining the earlier division into constraints of automation and affordances of augmentation.

5. Findings

In this section, we present our analysis of respondents' opinions on constraints related to the automation of the judging process in artistic gymnastics due to the implementation of an AI-powered system and affordances related to augmented human-AI judgment. The findings are organized along the three types of tasks: mechanical, thinking, and feeling tasks. An overview of our findings is presented in Table 3. From the voluminous interview data, we have selected the most comprehensive, interesting, and valuable material to support the presentation below.

Table 3. Constraints of automation and affordances of augmentation of judging tasks

Constraints of Automation			Affordances of Augmentation		
Mechanical	**Thinking**	**Feeling**	**Mechanical**	**Thinking**	**Feeling**
Mechanization of judgment Exactness of judgment Back-up for complete system break-down	Lack of precision for codifying rules Variety in rules Update of rules Athlete training “what”	Evaluate artistry Equalization of the athletes' performances and the loss of their personal style Capture human emotions Lack of human communication Human redundancy	Generating exact measures Speeding up the score generation and provisioning Harmonized and fault-less	Accurate scoring Judges oversight and competition Athlete training “how” Training to work with the system	Fair and objective judging Providing explanations Avoiding disputes

(Source: Authors' own creation)

5.1 Constraints of AI-Automation

We have identified the main constraints to automation of the judges' work by AI-powered judging system expressed and discussed by the interviewees. In this section, we present the constraints of automation of judging respectively in terms of mechanical, thinking, and feeling tasks.

5.1.1. Constraints related to Mechanical Tasks

Managing the technical infrastructure in gymnastics competitions is the responsibility of a dedicated technical committee. As part of this infrastructure, the judges interact mostly with a computer (i.e., to input their scores). In the incumbent human judging system, computers are only facilitating equipment. However, an AI-powered system that automates judging, according to the interviewees' opinions, might evoke some potential problems related to too *mechanization*, undesired *exactness* of judgment, and system's *break-down*.

According to the interviewees, a full automation of the judging process is impossible as judgment in artistic gymnastics is too complicated and requires the presence of human judges. The potential autonomous judgment of AI would make the sports too mechanical.

> *AI makes its forecast based on X-Y dimension and it will use some data program but actually gymnastics is extremely complicated. It's impossible for AI to give the final score. I think that anything used by the computer to make the judgment leads to low quality judging and low quality gymnastics. Because judging is the main feature of gymnastics. It is very important. Talking about using the computer to replace the judges to provide full judging by the system is impossible. It will make the sport become more like a machine. It can measure only angles, only time. (Harry, judge)*

Besides, the system will provide too high exactness of judgment. With the incumbent human judging, there is a reasonable balance between approximate judgment and the imperfection in athletes' execution. However, human athletes cannot match the exactness of AI-based judging with such level of perfection in their performances.

My worry is that the system is too perfect. It's a big difference: what a human eye sees is one thing, but what the machine sees is more accurate. Right now, we're humans. Gymnasts are humans. We as judges note certain deductions, certain angular deductions. Sometimes 45 degrees is very difficult to recognize for a human eye. But if a camera sees '44.9 degrees,' it does not accept the exercise; it makes a deduction. But for a human eye, the normal eye, it may pass. The gymnasts will be mad at the judgment with the machines because it's going to catch every single mistake they make. It will ask perfection of the gymnasts. Too much perfection. [...] In gymnastics if we eliminate all the judges and we only have this system, sport will be way different, it will be way lower quality, because the machine is too perfect. (Edward, judge)

With computers having a facilitating in the incumbent human judging system, in case of breaking down, judges can continue the competition with some more manual work (i.e., calculations). However, in case of a full automation of the judging process and a break-down of an AI-powered system, the competition would be severely interrupted, which is a major concern:

I would not allow the system to judge alone. In case it breaks down, there is no more competition. It is also very bad publicity. Imagine if it happens at the world championships, the system breaks down, the public will no longer watch it. (Lilly, judge)

I think someone always has to be there backing the system up. [...] it's new and you can't fully rely on it. If it's used at, for example, the world championship, it has to be secure. (Thomas, gymnast)

5.1.2. Constraints related to Thinking Tasks

Constraints related to the implementation of the AI-powered system for automation are linked to the complexity and variety of the rules of gymnastics (i.e., Code of Points), and the deficiency of the system to support athletes' training. The rules of judging in artistic gymnastics are very complicated and not formulated precisely enough to transmit into code for AI:

The rules of gymnastics are so complicated, it pushes the limits of human processing. Yet, they have not been able to develop the technology to judge a gymnastics routine. And they're realizing it's more complicated than thought and in part because there are flaws within the *rules [, as they have no…] objective measures. The FIG is trying to clarify the rules to make them more compatible with the information system. (Katarina, judge)*

There's a [score description] like "slightly" or something like that. We have to define what is "slightly" or what is "larger". So, we have to define it between 180-250 degrees or maybe 75 degrees or something. (Joona, Fujitsu)

In addition, the rules vary across regions and countries and are expected to evolve over time. So, the system must adapt and evolve continuously:

Looking at the grassroots and developmental levels, I don't see the [AI-powered] system at this point making all of the adaptations for all of the rules for all of the countries, for all of the levels. In the US, there are more than 17 variations of the rules, for example, and it's different for every country. (Katarina, judge)

There are so many variations and skills, they are different now from what they were ten years ago, even if these are the same skills, and the rules evolve, so the system must evolve at the same time. The equipment improves, and the gymnasts are better, they are faster and stronger, and they do it differently than a couple of years ago. So, the system has to evolve simultaneously. (Steven, technical director)

Stakeholders also pointed to potential constraints regarding the provision of the training support to the athletes by the system. While the system could help detect possible flaws during athlete training, the recommendation of *how* to improve their performance should be provided simultaneously, which the system is not capable of. Therefore, the training process cannot be fully automated only by the use of an AI system:

It is like saying to a gymnast, 'You have to jump higher', the system will also say that 'You jump at 1m20, that is too low', the system won't say how you have to jump higher. As a trainer, you need to do a practical translation. […] So I don't find it a big added value for training. (Sofia, judge)

The coach might think that some skill in the gymnast's performance looks good, but then s/he can look at the AI, and it could say what was actually wrong. The coach can check the faults *with the AI and know what is needed to be corrected by the gymnast. (Steven, technical director)*

5.1.3. Constraints related to Feeling Tasks

One of the main constraints related to the automation of the Feeling tasks mentioned by the interviewees is related to the inability of the system to judge the artistry and creativity of the athletes' performance. According to the stakeholders, these are an integral and crucial part of artistic gymnastics, which is why it is called "artistic" gymnastics. This specifically involves facial expressions, the choice of music, costumes, etc. At this point, AI cannot evaluate this part of gymnasts' performance:

> *There's an artistic component that can never be evaluated, in my opinion, by an artificial intelligence. It's like listening to a computer, doing a piece of Mozart's music, it can sound perfect, but it won't have a soul. It won't have emotive aspects. (Josh, judge)*
>
> *Part of the artistic deductions is the expression and the emotion, the effort put into the performance. And the technology that I've heard about at the moment would not be able to differentiate between those. It's about technical angles. And I think the interpretation of artistic expression is subjective. Some people might like it, some people might not like it. Everyone has different opinions. So, technology can't evaluate artistry, not yet, but who knows, maybe in the future. (Abby, judge)*
>
> *I don't think that it can replace. It's so subjective, that the AI can't replace the artistic part. How will you measure this? It's not possible. The computer will not see the faces of the gymnasts. This is the most important – the face, eyes, and smile – how can the computer analyze this? And then the music? Some people love the music, others don't love the music. (Lilly, judge)*
>
> *There is still the human aspect, especially on floor exercise, where it's important that you bring over emotion in your choreography to the judges and the audience. A computer cannot perceive and evaluate that. (Damian, gymnast)*

Therefore, stakeholders are afraid that the inability of an AI-powered system to evaluate the artistry of an athlete's routine and the subsequent elimination of the artistic part of the performance will lead to the equalization of all athletes and the loss of their personal style: "*In the end, we will have every exercise in the same way. Because they will try to make the exercise on the computer, but the personal style will be lost in the end." (Lilly, judge)*

Another crucial concern of the stakeholders was related to the inability of an AI-powered automated system to provide human (non-verbal) communication. Even though human communication is not an obvious part of judging work, especially during the actual competitions, it is very important. While gymnasts and judges are officially not allowed to communicate with each other during the actual competitions, non-verbal communication between them still happens. For instance, there is a non-verbal mutual greeting between the gymnast and the panel of judges, and judges' facial expressions after the routine, may reveal approval or support. Also, during podium training, judges don't only give advice and guide athletes on improving their routine, but also motivate and support them, encouraging them to obtain better results. The same is going on in between the competitions, and during the interactions of the judges with the national teams at various training camps.

According to the interviewees, as this AI-powered system on itself is not able to provide the human communication that gymnasts are used to, it would significantly impact the sports, making it more trivial, boring and uncomfortable for the athletes.

> *Is it going to be the best thing for the gymnasts? I think this system is going to take less into account the humanness of the gymnasts. That's what I believe. The human aspect is definitely one of the things that I worry about. (Sarah, judge)*
>
> *I'm not quite sure how the athletes will feel. Because I think that when an athlete does a good exercise and looks over to present to the judge and sees the reaction of the judge, I think that's something that is a human emotion that gives that athlete a good feeling [...]. Or when it's not so good and the judge has a sympathetic look even though the routine was not good, maybe the athlete still knows that there's someone who is cheering about the performance. You always look at the judges, right? So, the judges react to everything, and the gymnasts sometimes also react in turn. And if you're not making sure you have enough energy to do something and judges cheering you on, it can give you adrenaline and you can successfully*

do something. Well, I'm not sure if artificial intelligence will be able to provide that type of feedback to the athlete. (Charlie, judge)

I like the human aspect as well. Gymnasts standing in front of the computer and saying "Hi, I'm starting my exercise. That's kind of weird for me. (Nick, judge)

But honestly, to take the human aspect out of judging is not really a good idea. I think the system is lacking warmth. It's going to be a judge that smiles at you as you go, and then you will have an ugly camera staring at you. (Sarah, judge)

A similar constraint related to the automation of the judging process in artistic gymnastics is that the human judges could be replaced by AI one day. Judges are an important human capital of each competition as well as of the world community of artistic gymnastics and international and local federations. Their contribution to the development of the sport is crucial and can not be replaced by AI, according to the stakeholders. However, the fear that a new AI-powered judging system evoked in judges by the system negatively impacts their well-being, confidence in their future careers, and the future of sports.

There's another point that you have to think about- the people that are now working here as judges-they are the main judges of their countries. So, in their countries at home, they are doing the judges' courses for other people, and they are working in the Federations. If you cut them, all the Federations will lose their best people. Why should I be a judge if I can't do any more judging? [...]If it can provide the final score in the future, then you don't need the judges anymore. This is the idea. But, of course, I would prefer that it will not replace human judges because I like to be the judge and I would like to continue to do what I do now. (Lilly, judge)

Maybe in ten years, we will not have judges. But I'm a judge I love to do what I do, just afraid not to be needed anymore. Of course, I want it keep my job. I love what I do. (Bella, judge)

I don't believe that this system can replace human judges totally. No, it's not possible.[...] Besides, I would want personally to continue judging and not be replaced by some … robot. (Felicity, judge)

In the end, if the system proves to be more valuable than humans, I think that then it would be possible for the system to replace human judges [...] I would not use it personally to replace judges. I would be concerned if I felt like artificial technology was going to replace the judges entirely. I would be concerned about the future of the sport. (Abby, judge)

5.2 Affordances of AI-Augmentation

Overall, the stakeholders positively see the opportunity of a cooperative judgment with an AI-powered system:

The goal is to be able to help the judges in cases where better accuracy is needed, help the judges' education, help the coaches and the athletes with the training, improve the safety, and most importantly, hopefully, speed up the time of judging. So, when there are doubts, the answers are already available. (Simon, FIG)

We should take all the benefits of it. We should not perceive this system as an enemy. We should take it as a help for the judges. In those moments when it's so difficult for the human eye to see something, that machine should help us, so we should combine it together. For me, it's the best solution. (Felicity, judge)

I think that the combination of human judges and the system could be a good thing and could help the judges. (Abby, judge)

In our analysis, we identified many affordances across the three types of the judges' tasks: mechanical, thinking, and feeling, as evidenced in the subsections below.

5.2.1. Affordances related to Mechanical Tasks

The affordances of implementing AI into the mechanical tasks of the judges for joint work with human judges involve *generating exact measures* (i.e., in terms of duration or height), *speeding up the score generation and provisioning*, and doing so in a *harmonized and faultless way*.

The judges acknowledge that their human capabilities may be limited in terms of the (sensory) observation and (cognitive) processing of duration and height elements of an exercise. The technical capabilities of the AI-powered system could surpass the human ones in this regard. Hence, an affordance of the system lies in generating an exact measurement of duration and height. The following quotes illustrates this affordance:

> *I think this system has to be used […] to help the judges. We, judges, note certain deductions, certain angular deductions. You don't do 0.5, you may give 0.3 maybe. It's difficult if you sum up all the elements the gymnasts do. So, the system would be helpful and useful for these precise measurements. (Edward, judge).*
>
> *On the Rings [apparatus], you can use this system to measure the time[…] and it will be objective. (Harry, judge)*
>
> *The AI can be more precise, if it can sense […] the length, then it would be good to inform the judges. (Kyle, coach)*
>
> *It could be used for everything where deductions for height and distance can be taken, for us as judges that is always a rough estimation, "what is the height?"[as] there is no reference. (Emily, judge)*
>
> *You can use a third line on the Floor and for the Vault because sometimes the mats move a little bit, if you could put sensors in and sensors could pick whether someone's foot went over the line, then that would be fairer than using the eye. (Abby, judge)*

Other affordances relate to the score generation and provision happening faster. Indeed, the actual provision of the scores by judges involves many time-consuming mechanical tasks (see Table 1 and Appendix 1). The related affordances of implementing an AI-powered system do not only relate to speeding up the score generation and provisioning, but also lie in overcoming the lack of harmonization and potential errors (in symbol notation). Indeed, with an AI-powered system there would be no need to keep first the written report of the scores and deductions of each judge in a complicated and potentially problematic symbol notation, and then input them in the computer to generate the average score. Instead, the AI-powered system could generate and provide all scores and deductions directly on the screen, as the following quotes illustrate:

The human eye and human brain can't work as fast and accurate as the system. There are too many decisions to be taken, so for a human brain it is not possible to do it. (Nick, judge)

If the [AI-powered system] looks at the element and it gives you the deductions, then, yes, it can help to be faster in your judgment. (Lilly, judge)

5.2.2. Affordances related to Thinking Tasks

In our study, several affordances were identified related to the augmentation of the thinking tasks of the judges by AI. The main thinking tasks of the judges relate to the actual evaluation of the athletes' performance at the competitions and the provisioning of the scores and deductions, in keeping with the Code of Points. However, due to the initial complexity of the evaluation process sometimes the human cognitive capabilities of the judges are not enough for the evaluation of the athletes' performance with high enough accuracy:

It's about how the rules are created: in one second you have to make maybe 8 to 10 decisions, and you have to evaluate 1, 3, 5 and this is almost impossible because it's all at the same time [...] We support the development of this system because we need to use it because the human eye and human brain can't work so fast and accurate as the system like this. Simultaneously, there are too many decisions to be taken, so for a human brain it is not possible to do. (Norman, judge)

The cooperative work of the judges and AI-powered judging system for the evaluation of athletes' performances during the competitions might create such affordances as a *higher accuracy in generating the scores, the potential the system offers for judges' oversight and competition, and training affordances for athletes (explaining the "what").*

In the opinions of the stakeholders, complementing the judges' work with an AI-powered system during the competition can help to increase the accuracy of the evaluation of the elements of the gymnasts' routine, in particular for evaluating angles due to the advanced capabilities of AI:

As far as the human's ability to accurately perceive the angles, it is very difficult to differentiate between, for example, 44 and 45 degrees. How well can you do that from where *you're sitting? The AI system can be more accurate as far as that's concerned and will bring about more accurate scores. (Katarina, judge)*

I think that most judges want the right scores for the gymnasts. So, for using it for angles to say "Yes, that was completed within 10 degrees, you should not take the deduction, or it passed 10 degrees, yes, you should deduct" - that's helpful. (Abby, judge)

In addition, an AI-powered system could help with maintaining oversight of the judges. Currently, there is a control and monitoring system of the judges in place, which aims to identify cheating or unfair judging behavior. For instance, there might be suspicion if some judges provide a score critically different (lower or higher) from others. In the opinion of the stakeholders, the AI-powered system could help in detecting suspicious behavior and may even lead to positive competition among judges:

In my opinion, [an AI-powered system] can be used [...] post-competition, as an oversight mechanism in conjunction with the judges' evaluation program. […] If a judge knows that their score is ultimately going to be compared to the [AI-generated] score, then there will be a desire to try and make your score closer to what you anticipate the [AI-powered] system would award. […] In my fantasy, I would see judges having their competitions, looking at who can approximate the [AI-powered] system the closest. The judges take pride when they nail the score, if you come up with the exact same, it's like bingo, especially if it is across the panel. We pride ourselves when we get it right. (Katarina, judge)

As discussed above, an AI-powered system can not execute autonomous work on improving the training process of the athletes as it provides only the information on "what" should be improved in the athletes' routine but cannot explain "how" some elements should be improved. However, according to the opinion of the interviewees, augmented by the explanation of the human experts (coaches and judges) this system could become a good training instrument for the athletes, including for preventing injuries:

If FIG would provide this system as a training tool, especially for the gymnasts and coaches in their gym- that is where most of the action occurs, judges would be more willing to accept this system. For me personally, that's more critical than even the judging phase. Of course, this system is primarily a judging tool but I think that it's more important for the gymnasts and coaches for training in the gym. (Simon, FIG)

What, I think, could be the most valuable thing is the help that this system can provide in training, the gym, what the gymnasts do on a weekly basis, assisting the level of improvement or skill, or injury prevention, etc. [...] If it's (this system) showing me rotation and twisting, splitting, time in the air that the gymnast has to complete– all of these are like an amazing teaching aid. And this technical aid for the coaches, which I think would be really brilliant. [...] I think that the gymnasts can then get a better understanding of their technic, of the errors, and maybe also of how they're getting injured. So that to me is the most valuable thing. I see it as a key to perfecting training. (Sarah, judge)

Our analysis also revealed that the cooperative work between AI and judges for executing of the thinking tasks of the judges would require not only careful and step-by-step incorporation of the AI-powered system into judges' work and delegation of corresponding tasks of the judges to the system, but there is also a need for studying and training on how to work with the system (for system adoption and use):

We all have to know what the system is doing to have a really good opinion about it and be able to work with it. (Nick, judge)

You need to know how it works. If you have to work as a judge, then you need to know how to use this system. And you need to know what it can tell you, what information it can give you and how it gives you the information, and also how quickly it could give you the information. (Abby, judge)

The first step in my opinion is to make a pilot of the system and to do open training camps where people test it and provide their feedback on it. (Simon, FIG)

5.2.3. Affordances related to Feeling Tasks

Finally, a set of affordances was identified related to AI augmentation of the feeling tasks of judges. These tasks can be decomposed into judging and communication tasks. The affordances related to judging itself is that an AI-based system can be more *fair and objective*. For communication, affordances relate to *providing explanations*, which may *avoid disputes*. We discuss and illustrate each affordance in turn.

A long-time complaint in gymnastics is a perceived lack of objectivity and fairness of human judges, which can follow from various factors, such as emotions, personal preferences, initial expectations, familiarity with the routine or athlete, prejudice for or against a particular country:

> *Some judges think that they need to give higher scores to the athletes from the leading countries, and sometimes they even purposefully don't notice some mistakes of the athletes from these countries and give higher scores. It happens often that two athletes perform almost at the same level but the higher score is given to the one who is from the eminent country. (Ulla, judge)*

The hopes of the stakeholders are that with the help of an AI-powered system, judging may be more neutral and impartial, eliminating human biases and subjectivity:

> *The immediate impact that [an AI-powered system] can potentially have is to help bring about more unbiased judgments from the judges on the floor. […] And then the adjustment of the humans to what that ideal objective accurate score would be, will happen. And where those two meet, that's going to be super. (Katarina, judge)*
>
> *AI doesn't care which country you're from. It evaluates the technical side of the performance. Judges can hear very often from the coaches that we've been biased with their athletes, and if the routine is evaluated by the system, who can you blame for low scores? Nobody. Because AI is unbiased. It's objective. (Ulla, judge)*

A final set of affordances and constraints relates to communication tasks. As judges usually don't provide any explanation during the competitions, often it becomes unclear for the athletes why some scores were given or deductions were made. So, if the athletes don't agree with a score, the coaches can submit an inquiry to the judges asking to revise the score. However, these situations can become worse, as the revised score might be even lower than the original one, and again without any explanation from the judges. These situations can then end up with conflicts and disputes between the teams and the judges:

> *People are emotional. Coaches are often even more emotional than athletes. Very often it happens that an athlete works, trains well, and at the competition is also trying hard – of course. And for the coach, it may seem like the athlete did everything great! And then suddenly*

the judges punish and make deductions... and it happens that the coach comes, yelling emotionally, [...] "What are you judging here?! How can you do it?! My gymnast just did a great job!" [...] Very often, later on, they calm down and often come back, asking for forgiveness. (Ulla, judge)

According to several stakeholders, an AI-powered system has a good potential for providing explanations and as such resolving or even avoiding possible conflicts:

The system can always show where the deductions came from, why some skills were recognized, and why they were not recognized. (Steven, technical director)

When the system can provide some explanation or even a printed list of all deductions and scores, that would be great! Then it will be clear for everybody – for both coaches and gymnasts – how the judgment was done, and everybody will understand everything. (Ulla, judge)

There will be less discussion, and everyone will know for sure that the score is correct. Sometimes it is like "She has won by 1 tenth, if you had put one more mark, she wouldn't have won," this is some critique we get a lot. The technology could reduce those discussions, people will have more faith in it, and it will be more clear, so everyone will agree that the technology is correct and cannot be doubted. (Tracy, judge)

In addition, the provision of an explanation could potentially decrease the number of inquiries:

The system could help me, if for example, we are wondering whether it was a full turn or not. So, if the system could help me with that, then there will be no need for the inquires. (Felicity, judge)

What I heard is that it's very useful because there were a couple of inquiries especially on rings, that the human eye said "no", but when they saw the angle in 3 dimensions, there was very little difference. The human eye can't detect it. But the system did, so the inquiry was accepted. (Edward, judge)

It shows the angles and how many degrees the difference from the perfect angle was.

And that could be helpful when there's an inquiry for example. (Nick, judge)

6. Discussion

We discuss our findings first from a socio-technical perspective and then we discuss implications for research on AI automation and augmentation.

6.1 AI Affordances and Constraints from a Socio-Technical Perspective

In order to develop an understanding of whether and how different stakeholders are willing to accept and use AI-powered systems, it is important to study the complex socio-technical and institutional context involved (Leonardi, 2023; Sarker *et al.*, 2019; Seidel *et al.*, 2018). A way to do so, is to consider the affordances and constraints, namely how the technology may help or hinder goal-oriented behaviour (Strong *et al.*, 2014). We add to prior work that has identified affordances and constraints of AI in specific use cases, such as predictive maintenance (Keller *et al.*, 2019), predictive policing (Godé *et al.*, 2020), conversational agents (Stoeckli *et al.*, 2019; Waizenegger *et al.*, 2020), and auditing applications (Yang *et al.*, 2023). We also add to the emerging work that applies the affordance lens to situations in which technology has not been fully implemented or not even been experimented with (Du *et al.*, 2019).

We find that the constraints of automation and the affordances of augmentation are interdependent. For instance, by overcoming a hurdle in terms of providing precise rules (constraint of automation), an AI-powered system can provide accurate scoring (affordance of augmentation). Similarly, some of the constraints of automation are complementary to affordances of augmentation (e.g., athlete training the "what" and "how"). As Volkoff and Strong stated (2013, p. 828), it is therefore important to "appropriate enabling, stimulating, and releasing conditions" for the actualization of affordances. In addition, we can distinguish first and second-order affordances (Leidner *et al.*, 2018; Waizenegger *et al.*, 2020). For instance, implementing AI for thinking tasks can make evaluations more accurate, and a secondary affordance is the ability to use it for judges oversight.

Our findings attest that the perception and opinions of the key stakeholders were not systematically taken into consideration in the process of designing, developing, and testing an AI-powered system. This might lead to a low level of technology adoption by the key stakeholders (Sarker *et al.*, 2019; Yu *et al.*, 2023). In our study, we observed that despite all the positive sides and benefits that AI-powered judging can bring to competitive sports, many key stakeholders demonstrate negative perceptions, non-acceptance, and little readiness to cooperate with this technology. Since organizational institutions, technologies, and human workers are ontologically indistinguishable and intertwined as agents, any affordances or constraints provided by the technology will impact and entail changes in institutional values (Leonardi, 2023). In our case study, such constraints of the system as the inability to evaluate the athlete's artistry, being one of the core values of artistic gymnastics, can be perceived as highly problematic. Hence, we assert that involving the key stakeholders more intensively in the process of such systems' implementation and increasing their awareness of the capabilities and its role in future evaluation processes can enable the "collaborative optimization, fit and harmony" of institutional, social, and technical elements (Sarker *et al.*, 2019, p. 704). To do so, one must proactively give stakeholders a critical voice in how that system will be used (Tong *et al.*, 2021). This may result in the successful tailoring and integration of an AI-powered system, while maintaining well-being and satisfaction of stakeholders (Sarker *et al.*, 2019).

6.2 Implications for Automation and Augmentation of Work Task Types

In this section, we discuss what implications can be derived from our findings in terms of the two main approaches to implement AI: automation and augmentation. Because each approach involves distinct and sometimes even conflicting measures for implementation (Benbya *et al.*, 2021), a nuanced understanding of why one approach may be preferred over the other is important. We keep with the notion that the appropriate approach to AI implementation in the organization depends on the task characteristics (Seidel *et al.*, 2019). In our study, we have applied the division of three task types (mechanical, thinking, feeling) to the judges' work in artistic gymnastics (Huang and Rust, 2018).

Our study aligns with previous research finding that there are constraints to the automation of thinking and feeling tasks so that thinking tasks can be automated only to some extent, while feeling tasks cannot be automated at all. Establishing an autonomous system for implementing the thinking tasks is difficult because of issues related to codifying the rules, especially given the latter may vary and evolve. For feeling tasks, prior research indicates that despite the advanced stage of the development of feeling AI (Huang and Rust, 2024), its capabilities are still limited at this point in time. Indeed, our findings provide granular insight by showing that AI's lack of sensing human emotions and communicating non-verbally, as well as its inability to evaluate the creativity and artistry of athletes' performances, makes it impossible to automate the feeling tasks of the judges. While AI has superior computational "hard skills," interpersonal "soft skills" are still paramount in sports judging (Luo *et al.*, 2021).

However, in contrast to previous studies, stating that mechanical tasks can be fully automated and implemented by AI with limited adaptability and learning (Jia *et al.*, 2024; Raisch and Krakowski, 2021), our study shows that in the case when all three types of tasks are closely intertwined, there are constraints to the automation of mechanical tasks as well. Stakeholders state that automation can actually engender changes that go against the values they care deeply about (van den Broek *et al.*, 2021). Constraints such as too high mechanization of judgment, as well as too high exactness of judgment that contradicts the human imperfection of athletes' performances would negatively impact the whole sport and its future development. Similarly, the system was envisioned to provide training possibilities for athletes. For such functionality, our findings point out that while AI can be adequate in communicating the know-what, conveying the know-how is the more difficult part (Lebovitz *et al.*, 2021).

In addition, we found that automation might evoke negative perceptions of the stakeholders and lower their productivity due to the lack of acceptance and understanding of AI (Baer *et al.*, 2022; Crawford and Whittaker, 2016). Being concerned about the automation of the evaluation process of artistic gymnastics, re-structuring and re-consideration of the whole sport due to AI use in it as well as potential job losses, the interviewees demonstrated a quite high level of anxiety and uncertainty about their professional careers (Daugherty and Wilson, 2018; Yu *et al.*, 2023). This in turn negatively impacts their acceptance of the technology and the process of its effective implementation in the evaluation process. Our findings point out that although human and technological agents are closely interconnected and do not exist as individual elements in institutional practices (Leonardi, 2023), the negative perception of a new AI-powered system related to core institutional values, undermines an effective collaboration.

Therefore, augmentation can be a better approach to the implementation of AI into the judges' work in artistic gymnastics characterized by an intense interaction between AI and human judges for a more effective performance of the tasks (Rai *et al.*, 2019). Joint work between the judges and AI can provide such benefits as higher accuracy and objectivity of the judging process, speed up the evaluation process, and help in resolving potential conflicts and misunderstandings that from time to time arise at the competition between the judges and the teams. An augmented approach would enable AI to take over some judges' mechanical and repetitive tasks and partly the thinking tasks, while the overall monitoring over the judging process, control of the final decision-making, and overall accountability for the result would still remain with human experts (Berente *et al.*, 2021; Seidel *et al.*, 2019; Wilson *et al.*, 2017). In this way, AI can be used for real-time processing of judging data, yet leaving the provision of the final scores to humans (Dellermann, Ebel, *et al.*, 2019). Put differently, judges' tasks can be augmented by AI, providing an accurate and precise evaluation and AI-powered judging system decision-making can be augmented by human capabilities (Baer *et al.*, 2022).

In terms of athletes' training, the information about the know-what to change in the routine based on the system's recommendation will be valuable for the coaches, but explaining the know-how remains a human capability (Lebovitz *et al.*, 2021). We also observed secondary possibilities, which are a consequence of digitalizing the process. Indeed, evaluating the gymnasts is the primary goal of an AI-based system, yet based on the data generated it could also allow developing a system for evaluating the judges and for personalizing training for judges (Luo *et al.*, 2021). It would enable continuous mutual learning and intense knowledge exchange between humans and AI (van den Broek *et al.*, 2021) when human judges can learn from AI how to provide a better quality evaluation process and AI can learn from humans' empathy, sensitivity, creativity and ethics (Daugherty and Wilson, 2018).

However, augmented human-AI judging will also require re-consideration of the judges' tasks and work as well as the performance of other stakeholders. A new joint human-AI evaluation in sports would require doing new things by the teams and doing things differently by the judges (Daugherty and Wilson, 2018). The AI implementation would require additional training for the judges on how to use the system and how to effectively collaborate with it (Daugherty and Wilson, 2018), while for the athletes to adjust to new judging criteria of joining human-machine evaluation that can become more precise than previously.

In conclusion, the granular approach taken in this study, allowed us to observe an intricate dance between constraints and affordances of two approaches to AI implementation across task types.

Indeed, constraints to automation and affordances to augmentation were found in each of the task types, which suggests that in the end, continuous collaboration between AI and experts can become the modus operandi (van den Broek *et al.*, 2021; Daugherty and Wilson, 2018).

7. Conclusion

In this paper, we have identified affordances and constraints of approaches towards AI implementation across task types, based on perceptions of multiple stakeholders. We hope that examining AI in this granular and comprehensive manner may provide a fruitful avenue to debate societal and organizational implications of using AI, such that the technology can realize its transformative potential (von Krogh, 2018; Raisch and Krakowski, 2021).

Our study also has important practical implications. Indeed, implementing AI-powered systems in inappropriate ways may lead to negative-reinforcement cycles and trigger unintended consequences for an organization or society at large (Ågerfalk *et al.*, 2022; Raisch and Krakowski, 2021). Indeed, if an AI-powered system is found to be functionally limited or experienced as less adequate for tasks previously executed by human experts, it is likely to erode positive perceptions and support (Raisch and Krakowski, 2021). Moreover, insufficient (financial) resources for AI implementation might lead to even bigger failures and losses (Raisch and Krakowski, 2021). Our paper suggests that, rather than making decision to automate or augment jobs, organizations should decompose jobs into task types but don't even stop there. Indeed, at that level lists of constraints and affordances can be identified, which can lead to conscious and cautious paths of AI implementation. To test the applicability of our findings, we conducted a limited set of checks (Rosemann and Vessey, 2008). Indeed, we presented our research output to a judge and to a federation and our findings passed these checks, "indicating that it addresses a problem that is important for practitioners in an understandable manner" (Rinta-Kahila *et al.*, 2023, p. 1388).

A limitation of our paper is that it was based on a single case study in a specific context that is unique due to the specific stakeholders involved (Stremersch *et al.*, 2022). Such research has been criticized for providing limited generalizability (Stremersch *et al.*, 2022; Walsham, 2006). While single case studies are common in IS research (Lee and Baskerville, 2003, p. 231), we acknowledge that generalization in a statistical sense is impossible with our research design. However, we contend that one can generalize our study beyond its singular context because we grounded the framework from the case's empirical reality and corroborated it with an established framework on task types (Huang *et al.*, 2019). Moreover, going beyond the gold standard of generalizability, strengths of a context-specific study include high internal validity, high meaningfulness for various stakeholders, and high potential to fuel innovation (Stremersch *et al.*, 2022). Nevertheless, future research could examine various approaches towards AI implementation in other contexts.

References

Ågerfalk, P.J., Conboy, K., Crowston, K., Eriksson Lundström, J., Jarvenpaa, S.L., Ram, S. and Mikalef, P. (2022), "Artificial Intelligence in Information Systems: State of the Art and Research Roadmap", *Communications of the Association for Information Systems*, Vol. 50 No. 1, pp. 420–438.

Agrawal, A., Gans, J.S. and Goldfarb, A. (2017), "What to expect from artificial intelligence", *MIT Sloan Management Review*, Vol. 58 No. 3, pp. 23–26.

Asatiani, A., Malo, P., Nagbøl, P.R., Penttinen, E., Rinta-Kahila, T. and Salovaara, A. (2021), "Sociotechnical Envelopment of Artificial Intelligence: An Approach to Organizational Deployment of Inscrutable Artificial Intelligence Systems", *Journal of the Association for Information Systems*, Vol. 22 No. 2, pp. 325–352.

Autio, E., Nambisan, S., Thomas, L.D.W. and Wright, M. (2018), "Digital affordances, spatial affordances, and the genesis of entrepreneurial ecosystems", *Strategic Entrepreneurship Journal*, Vol. 12 No. 1, pp. 72–95.

Baer, I., Waardenburg, L. and Huysman, M. (2022), "What Are We Augmenting? A Multidisciplinary Analysis of AI- based Augmentation for the Future of Work", *International Conference on Information Systems*, pp. 1–17.

Balasubramanian, N., Ye, Y. and Xu, M. (2022), "Substituting Human Decision-Making With Machine Learning: Implications for Organizational Learning", *Academy of Management Review*, Vol. 47 No. 3, pp. 448–465.

Benbya, H., Pachidi, S. and Jarvenpaa, S.L. (2021), "Special issue editorial: Artificial intelligence in organizations: Implications for information systems research", *Journal of the Association for Information Systems*, Vol. 22 No. 2, pp. 281–303.

Berente, N., Gu, B., Recker, J. and Santhanam, R. (2021), "Managing Artificial Intelligence", *MIS Quarterly*, Vol. 45 No. 3, pp. 1433–1450.

Bogner, A., Littig, B. and Menz, W. (2018), "Generating qualitative data with experts and elites", *The SAGE Handbook of Qualitative Data Collection*, SAGE Publications Ltd, pp. 652–667.

van den Broek, E., Sergeeva, A. and Huysman, M. (2021), “When the machine meets the expert: An ethnography of developing AI for hiring”, *MIS Quarterly*, Vol. 45 No. 3, pp. 1557–1580.

Brooks, L.A. and Saveri, A. (2017), “Expanding Imagined Affordance with Futuretypes: Challenging Algorithmic Power with Collective 2040 Imagination”, *Proceedings of the 50th Hawaii International Conference on System Sciences (2017)*, pp. 1765–1774.

Brynjolfsson, E., Rock, D. and Syverson, C. (2017), *Artificial Intelligence and the Modern Productivity Paradox: A Clash of Expectations and Statistics*, *NBER Working Paper*, Vol. Publisher.

Constantinides, P., Monteiro, E. and Mathiassen, L. (2024), “Human-AI joint task performance: Learning from uncertainty in autonomous driving systems”, *Information and Organization*, Elsevier Ltd, Vol. 34 No. 2, p. 100502.

Crawford, K. and Whittaker, M. (2016), *The Social and Economic Implication of Artificial Intelligence Technologies in the Near-Term*.

Daugherty, P.R. and Wilson, H.J. (2018), *Human + Machine: Reimagining Work in the Age of AI*, Harvard Business Press, Boston, MA.

Davenport, B.Y.T.H. and Ronanki, R. (2018), “Artificial Intelligence for the Real World”, *Harvard Business Review*, No. 1, pp. 108–116.

Day, D. V, Gordon, S. and Fink, C. (2012), “The Sporting Life: Exploring Organizations through the Lens of Sport”, *The Academy of Management Annals*, pp. 1–37.

Dellermann, D., Calma, A., Lipusch, N., Weber, T., Weigel, S. and Ebel, P. (2019), “The Future of Human-AI Collaboration: A Taxonomy of Design Knowledge for Hybrid Intelligence Systems”, *Proceedings of the 52nd Hawaii International Conference on System Sciences*, available at:https://doi.org/10.24251/hicss.2019.034.

Dellermann, D., Ebel, P., Söllner, M. and Leimeister, J.M. (2019), “Hybrid Intelligence”, *Business Information Systems Engineering*, Vol. 61 No. 5, pp. 637–643.

Du, W. (Derek), Pan, S.L., Leidner, D.E. and Ying, W. (2019), “Affordances, experimentation and actualization of FinTech: A blockchain implementation study”, *Journal of Strategic Information Systems*, Elsevier, Vol. 28 No. 1, pp. 50–65.

Ebel, P., Söllner, M., Leimeister, J.M., Crowston, K. and de Vreede, G.J. (2021), “Hybrid intelligence in business networks”, *Electronic Markets,* Vol. 31 No. 2, pp. 313–318.

Emirbayer, M. and Mische, A. (1998), “What is agency?”, *The American Journal of Sociology,* Vol. 103 No. 4, pp. 962–1023.

Essén, A. and Värlander, S.W. (2019), “How Technology-Afforded Practices at the Micro-Level can Generate Change at the Field Level: Theorizing the Recursive Mechanism Actualized in Swedish Rheumatology 2000-2014”, *MIS Quartely*, Vol. 43 No. 4, pp. 1155–1176.

Felin, T., Kauffman, S., Mastrogiorgio, A. and Mastrogiorgio, M. (2016), “Factor markets, actors, and affordances”, *Industrial and Corporate Change,* Vol. 25 No. 1, pp. 133–147.

Fonti, F., Ross, J. and Aversa, P. (2023), “Using Sports Data to Advance Management Research : A Review and a Guide for Future Studies”, *Journal of Management*, Vol. 49 No. 1, pp. 325–362.

Fujiwara, H. and Ito, K. (2018), “ICT-based judging support system for artistic gymnastics and intended new world created through 3D sensing technology”, *Fujitsu Scientific and Technical Journal*, Vol. 54 No. 4, pp. 66–72.

Godé, C., Brion, S. and Bohas, A. (2020), “The Affordance-Actualization Process in a Predictive Policing Context: Insights from the French Military Police”, *Twenty-Eighth European Conference on Information Systems (ECIS2020)*.

Goebeler, L., Standaert, W. and Xiao, X. (2021), “Hybrid Sport Configurations: The Intertwining of the Physical and the Digital”, *Proceedings of the 54th Hawaii International Conference on System Sciences*.

Huang, M.-H. and Rust, R.T. (2024), “EXPRESS: The Caring Machine: Feeling AI for Customer Care”, *Journal of Marketing*, No. Forthcoming, available at:https://doi.org/10.1177/00222429231224748.

Huang, M.H., Rust, R. and Maksimovic, V. (2019), “The Feeling Economy: Managing in the Next Generation of Artificial Intelligence (AI)”, *California Management Review*, pp. 43–65.

Huang, M.H. and Rust, R.T. (2018), “Artificial Intelligence in Service”, *Journal of Service Research*, Vol. 21 No. 2, pp. 155–172.

Jarvenpaa, S.L. and Standaert, W. (2018), “Digital Probes as Opening Possibilities of Generativity”, *Journal of the Association for Information Systems*, Vol. 19 No. 10, pp. 982–1000.

Jia, N., Luo, X., Fang, Z. and Liao, C. (2024), "When and How Artificial Intelligence Augments Employee Creativity", *Academy of Management Journal*, Vol. In-Press, available at:https://doi.org/10.5465/amj.2022.0426.

Keller, R., Stohr, A., Fridgen, G., Lockl, J. and Rieger, A. (2019), "Affordance-experimentation-actualization theory in artificial intelligence research - A predictive maintenance story", *40th International Conference on Information Systems*, pp. 1–17.

von Krogh, G. (2018), "Artificial Intelligence in Organizations: New Opportunities for Phenomenon-Based Theorizing", *Academy of Management Discoveries*, Vol. 4 No. 4, pp. 404–409.

Lebovitz, S., Levina, N. and Lifshitz-Assaf, H. (2021), "Is AI Ground Truth Really True? The Dangers of Training and Evaluating AI Tools Based on Experts' Know-What", *MIS Quarterly*, Vol. 45 No. 3, pp. 1501–1525.

Lee, A.S. and Baskerville, R.L. (2003), "Generalizing Generalizability in Information Systems Research", *Information Systems Research*, Vol. 14 No. 3, available at:https://doi.org/10.1287/isre.14.3.221.16560.

Leidner, D.E., Gonzalez, E. and Koch, H. (2018), "An affordance perspective of enterprise social media and organizational socialization", *Journal of Strategic Information Systems*, Elsevier, Vol. 27 No. 2, pp. 117–138.

Leonardi, P. (2023), "Affordances and Agency: Toward the Clarifiation and Integration of Fractured Concepts", *MIS Quartely*, Vol. 47 No. 4, pp. ix–xx.

Luo, X., Qin, M.S., Fang, Z. and Qu, Z. (2021), "Artificial Intelligence Coaches for Sales Agents : Caveats and Solutions", *Journal of Marketing*, Vol. 85 No. 2, pp. 14–32.

Luo, X., Tong, S., Fang, Z. and Qu, Z. (2019a), "Frontiers: Machines vs. Humans: The Impact of Artificial Intelligence Chatbot Disclosure on Customer Purchases", *Marketing Science*, Vol. 38 No. 6, pp. 937–947.

Luo, X., Tong, S., Fang, Z. and Qu, Z. (2019b), "Frontiers: Machines vs. humans: The impact of artificial intelligence chatbot disclosure on customer purchases", *Marketing Science*, INFORMS Inst.for Operations Res.and the Management Sciences, Vol. 38 No. 6, pp. 937–947.

Marino, A., Aversa, P., Mesquita, L.F. and Anand, J. (2015), “Driving Performance Via Exploration In Changing Environments: Evidence From Formula One Racing”, *Organization Science*, Vol. 26 No. 4, pp. 1079–1100.

Nagy, P. and Neff, G. (2015), “Imagined Affordance: Reconstructing a Keyword for Communication Theory”, *Social Media and Society*, Vol. 1 No. 2, pp. 1–9.

Patton, M.Q. (2002), *Qualitative Research and Evaluation Methods (3rd Ed.), Sage Publications*, SAGE Publications, available at:https://doi.org/10.1177/1035719X0300300213.

Rai, A., Constantinides, P. and Sarker, S. (2019), “Next-Generation Digital Platforms: Toward Human-AI Hybrids”, *MIS Quarterly*, Vol. 43 No. 1.

Raisch, S. and Krakowski, S. (2021), “Artificial intelligence and management: The automation–augmentation paradox”, *Academy of Management Review*, Vol. 46 No. 1, pp. 192–210.

Rinta-Kahila, T., Penttinen, E., Salovaara, A., Soliman, W. and Ruissalo, J. (2023), “The Vicious Circles of Skill Erosion: A Case Study of Cognitive Automation”, *Journal of the Association for Information Systems*, Vol. 24 No. 5, pp. 1378–1412.

Rosemann, M. and Vessey, I. (2008), “Toward Improving the Relevance of Information Systems Research to Practice: The Role of Applicability Checks”, *MIS Quarterly*, Vol. 32 No. 1, pp. 1–22.

Sarker, S., Chatterjee, S., Xiao, X. and Elbanna, A. (2019), “The sociotechnical axis of cohesion for the IS discipline: Its historical legacy and its continued relevance”, *MIS Quarterly: Management Information Systems*, Vol. 43 No. 3, pp. 695–719.

Sarker, S., Xiao, X., Beaulieu, T. and Lee, A.S. (2018), “Learning from First-Generation Qualitative Approaches in the IS Discipline : An Evolutionary View and Some Implications for Authors and Evaluators (Part 2/2)”, *Journal of the Association for Information Systems*, Vol. 19 No. 9, pp. 909–923.

Seidel, S., Berente, N., Lindberg, A., Lyytinen, K. and Nickerson, J. V. (2018), “Autonomous tools and design”, *Communications of the ACM*, Vol. 62 No. 1, pp. 50–57.

Seidel, S., Berente, N., Lindberg, A., Lyytinen, K. and Nickerson, J. V. (2019), “Autonomous tools and design: A triple-loop approach to human-machine learning”, *Communications of the ACM*, Vol. 62 No. 1, pp. 50–57.

Song, X., Xu, B. and Zhao, Z. (2022), “Can people experience romantic love for artificial intelligence? An empirical study of intelligent assistants”, *Information and Management*, Vol. 59 No. 2, pp. 1–10.

Stoeckli, E., Dremel, C., Uebernickel, F. and Brenner, W. (2019), “How affordances of chatbots cross the chasm between social and traditional enterprise systems”, *Electronic Markets*, Electronic Markets, Vol. 30, pp. 369–403.

Stremersch, S., Gonzalez, J., Valenti, A. and Villanueva, J. (2022), “The value of context-specific studies for marketing”, *Journal of the Academy of Marketing Science*.

Strong, D.M., Volkoff, O., Johnson, S.A., Pelletier, L.R., Tulu, B., Bar-on, I., Trudel, J., *et al.* (2014), “A Theory of Organization-EHR Affordance Actualization”, *Journal of the Association for Information Systems*, Vol. 15 No. 2, pp. 53–85.

Tong, S., Jia, N., Xueming, L. and Fang, Z. (2021), “The Janus face of artificial intelligence feedback: Deployment versus disclosure effects on employee performance”, *Strategic Management Journal*, Vol. 42 No. 9, pp. 1600–1631.

Volkoff, O. and Strong, D.M. (2013), “Critical realism and affordances: Theorizing IT-associated organizational change processes”, *MIS Quarterly: Management Information Systems*, Vol. 37 No. 3, pp. 819–834.

Volkoff, O. and Strong, D.M. (2017), “Affordance theory and how to use it in is research”, *The Routledge Companion to Management Information Systems*, pp. 232–246.

Waizenegger, L., Seeber, I., Dawson, G. and Desouza, K. (2020), “Conversational Agents - Exploring Generative Mechanisms and Second-hand Effects of Actualized Technology Affordances”, *Proceedings of the 53rd Hawaii International Conference on System Sciences*, Vol. 3, available at:https://doi.org/10.24251/hicss.2020.636.

Walsham, G. (2006), “Doing interpretive research”, *European Journal of Information Systems*, Vol. 15 No. 3, pp. 320–330.

Wang, W., Gao, G. (Gordon) and Agarwal, R. (2024), “Friend or Foe? Teaming Between Artificial Intelligence and Workers with Variation in Experience”, *Management Science*, http://pubsonline.informs.org/journal/mnsc, No. Forthcoming, available at:https://doi.org/10.1287/mnsc.2021.00588.

Wilson, H.J., Daugherty, P.R. and Morini-Bianzino, N. (2017), “The Jobs That Artificial Intelligence Will Create”, *MIT Sloan Management Review*, Vol. 58 No. 4, pp. 14–16.

Xiao, X., Chian Tan, F. Ter, Lim, E.T.K., Henningsson, S., Vatrapu, R., Hedman, J., Tan, C.W., *et al.* (2017), “Sports Digitalization: An Overview and A Research Agenda”, *38th International Conference on Information Systems*, pp. 1–21.

Yang, J., Marrone, M. and Amrollahi, A. (2023), “What Makes AI Different? Exploring Affordances and Constraints - The Case of Auditing”, *European Conference on Information Systems*, p. 265.

Yin, R.K. (2014), *Case Study Research: Design and Methods*, 5th ed., SAGE Publications, available at:https://doi.org/10.2139/ssrn.1444863.

Yu, X., Xu, S. and Ashton, M. (2023), “Antecedents and outcomes of artificial intelligence adoption and application in the workplace: the socio-technical system theory perspective”, *Information Technology and People*, Vol. 36 No. 1, pp. 454–474.

Appendix 1. Detailed Description of the Judging Task across Types

(Source: Authors' own creation)

Mechanical tasks	Thinking tasks	Feeling tasks
1. Facilitating the efficient and smooth running of the competition including: a) following the correct order of the teams and gymnasts, b) monitoring the time of warm-up, the commencement, and duration of the exercise, c) monitoring and timely informing of any faults or violations. 2. Provisioning and monitoring of all technical aspects of the competition, such as: a) ensuring the efficient running of the apparatus, computers, and other technical equipment, b) giving audible and light signals notifying the gymnasts of the start and finish of the competitions/exercises, c) ensuring all deductions for overtime limit, line, and behavior faults are applied. 3. Controlling and monitoring the scoring system: a) recording all scores and deductions in symbol notation on score sheets, b) entering the final scores into the computer, c) submission of a written report with all detailed information.	1. Participation in discussion and development of organizational and planning strategies and objectives for judging at the scheduled meetings of judges. 2. Obtaining and updating knowledge of judging rules in order to maintain qualifications as an international judge. 3. Getting the updated special judging-related instructions valid for the current level of competition at the relative sessions of judges. 4. Attending podium training and provision of consultation and advice to athletes, coaches, and teams on changing/improving their routines to obtain better results and higher scores. 5. During the competition: accurate, consistent, and quick judging of the quality of the athlete's performance according to the competition standards, and applying deductions corresponding to general and specific apparatus execution faults made by the athletes. 6. Monitoring and coordinating the work and activities of other judges, depending on the rank of judges.	1. Evaluating the athlete's performance in terms of artistry performance and creativity. 2. Exemplifying nonpartisan, ethical and professional behavior including objective and fair judging and taking full personal responsibility for the scores. 3. Resolving conflicts with the coaches/gymnasts related to disputes about scores. 4. Establishing and maintaining professional relationships and communicating effectively with the teams (gymnasts, coaches, etc.). 5. Communicating effectively and professionally with colleagues and representatives of governing bodies. 6. Training, teaching, and assisting gymnasts and other judges. 7. Communicating with the public, audience, fans, and media.

Appendix 2. Example Interview Questions (Source: Authors' own creation)

- What is your history with artistic gymnastics?
- What are the different tasks of judges?
- How do you feel about the possible implementation of an AI-powered system for judging in gymnastics?
- Do you support using the system? Why (not)?
- How do you think an AI-powered system could be implemented in relation to judging in gymnastics?
- What would be possible advantages or disadvantages of an AI-powered system in gymnastics?
- Do you believe it is possible to completely replace humans for judging in gymnastics? Why (not)?
- <To judges> How do you see an AI-powered system affecting gymnasts and coaches?
- <To gymnasts and coaches> How do you see an AI-powered system affecting judges?
- How do you see an AI-powered system evolve in the short and long term?

Appendix 3. Coding Scheme (Source: Authors' own creation)

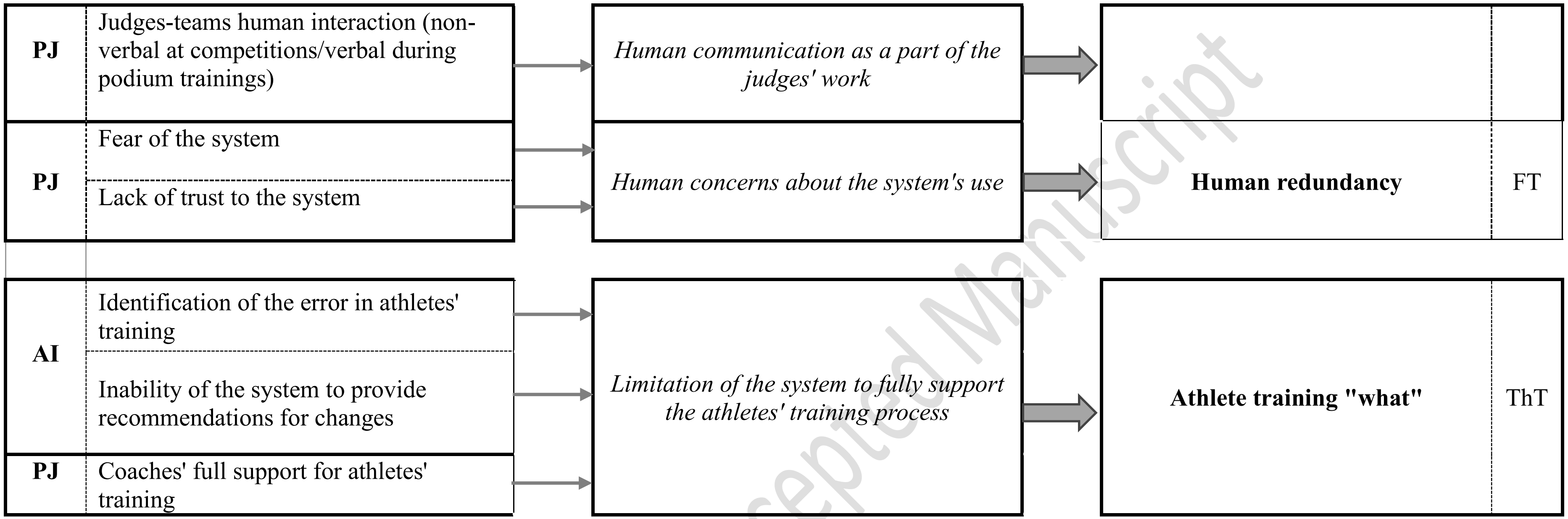
Glossary: **AI -** AI-powered system; **PJ -** Panel of judges;

MT - Mechanical tasks; **ThT** - Thinking tasks; **FT** - Feeling tasks

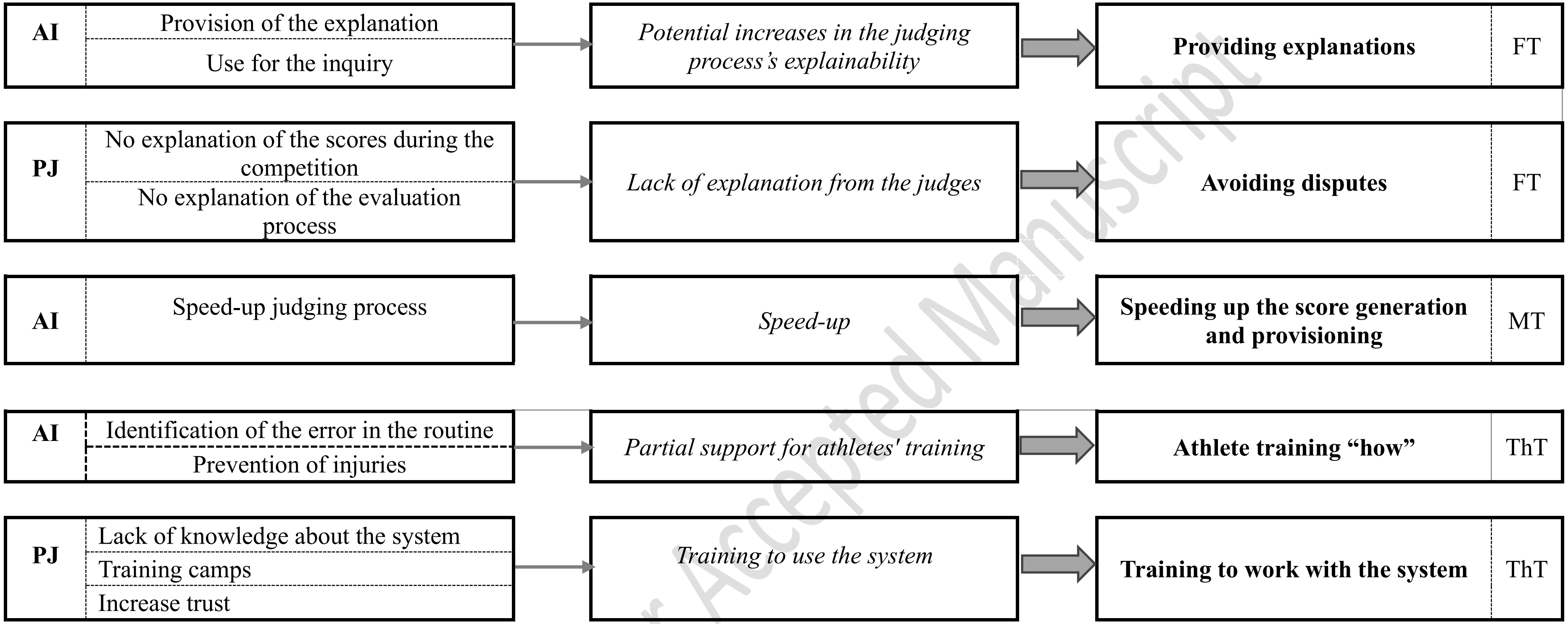
Constraints of Automation

	Open coding	Selective coding	Axial coding	
AI	System's technical capabilities	*Computation of judgment*	**Mechanization of judgment**	MT
AI	System's accuracy and precision	*Complexity of current judgment in artistic gymnastics*	**Mechanization of judgment**	MT
AI	Perfection/Exactness/Preciseness	*Too high perfection of the system*	**Exactness of judgment**	MT
PJ	Imperfection of judges equals imperfection of athletes	*Mismatch with imperfection of athletes*	**Exactness of judgment**	MT
AI	Systems' failure	*Incapability of the system for completely autonomous work*	**Back-up for complete system break-down**	MT
AI	Systems' error	*Incapability of the system for completely autonomous work*	**Back-up for complete system break-down**	MT
PJ	Interchangeability of judges	*The need of a human expert to back-up the system*	**Back-up for complete system break-down**	MT

PJ	Current Code of Points	*Inability of the system to adjust to the variety of rules and their constant changes*	**Lack of precision for codifying rules**	ThT
PJ	No objective measures in rules	*Changes in the rules*	**Variety in rules**	ThT
	No preciseness in rules	*The need for the clarity of the rules to make them more compatible with the AI*	**Update of rules**	ThT
	Many variations and skills	*Simultaneous changes of the system to the variety and changes of rules*		
AI	Inability to evaluate artistry	*Inability to evaluate artistry*	**Evaluate artistry**	FT
PJ	Emotion as a key component of gymnastics	*Inability to capture emotions*	**Capture human emotions**	FT
AI	Inability of the system to evaluate artistry	*High standardization of the judgment*	**Equalization of the athletes' performances and the loss of their personal style**	FT
AI	Subsequent elimination of the artistic part of the performance			
AI	Inability of the system for human interaction	*Lack of human interaction on the system's part*	**Lack of human communication**	FT


PJ
Judges-teams human interaction (non-verbal at competitions/verbal during podium trainings)
Human communication as a part of the judges' work
PJ
Fear of the system
Lack of trust to the system
Human concerns about the system's use
Human redundancy
FT
AI
Identification of the error in athletes' training
Inability of the system to provide recommendations for changes
PJ
Coaches' full support for athletes' training
Limitation of the system to fully support the athletes' training process
Athlete training "what"
ThT

Open coding		Selective coding	Axial coding	
		Affordances of augmentation		
AI	Accuracy of the system	*The higher accuracy of an AI-powered system*	**Generating exact measures**	MT
			Accurate scoring	ThT
PJ	Cognitive limitations of human brain	*The lack of accuracy of human judges*	**Harmonized and fault-less**	MT
	Human fatigue			
	Human error		**Judges oversight and competition**	ThT
	Judges' eyes			
AI	Objectivity of the system	*AI-powered system's greater objectivity and impartiality*	**Fair and objective judging**	FT
	Increased justice of judging			
	Athletes protection			
PJ	Subjectivity and prejudice of human judges:	*Subjectivity and prejudice of human judges*		
	Expectations of the judges			
	Familiarity with the routine			
	Friendship between the judges and coaches			
	The order of athletes' performance			
	Unofficial guidelines			
	Familiarity / expectations held for an athlete			


AI
Provision of the explanation
Use for the inquiry
Potential increases in the judging process's explainability
Providing explanations
FT
PJ
No explanation of the scores during the competition
No explanation of the evaluation process
Lack of explanation from the judges
Avoiding disputes
FT
AI
Speed-up judging process
Speed-up
Speeding up the score generation and provisioning
MT
AI
Identification of the error in the routine
Prevention of injuries
Partial support for athletes' training
Athlete training "how"
ThT
PJ
Lack of knowledge about the system
Training camps
Increase trust
Training to use the system
Training to work with the system
ThT

Appendix 4. Open Coding Examples (Source: Authors' own creation)

Constraints of Automation	
Perfection/Exactness/Preciseness	*"My worry is that the system is too perfect. It will ask perfection of the gymnasts. Too much perfection."*
System's failure	*"Because could you imagine if something happens to the system, we gonna ask the athlete to do the exercise again? No."*
System's error	*"There's always going to be maybe a chance for a system error because the error was made to programming the system."*
Many variations of rules	*"There are so many variations [..], they are different now from what they were ten years ago, even if these are the same skills, and the rules evolve, so the system must evolve at the same time."*
Inability of the system to evaluate the artistry	*"The computers don't understand what is artistic."*
Human interaction between judges and gymnasts	*"We're part of the competition, and there should be always a human aspect of judging at the competition."*
Inability of the system for human interaction	*"Gymnasts standing in front of a computer and saying, 'Hi, I'm starting my exercise.' That's kind of weird for me."*
Fear of the system	*"But of course you should understand that some atmosphere of among the international judges is fear. Is it so that in the future we're not needed?*"
Lack of trust in the system	*"You can trust it only it has been verified somehow. And it has to be done regularly.*"
Identification of the error in athletes' training	"*The coach might think that some skill in the gymnast's performance looks good, but then s/he can look at the AI, and it could say what was actually wrong."*
Inability of the system to provide recommendations for changes	"*It is like saying to a gymnast, 'You have to jump higher' [...] the system won't say how you have to jump higher. As a trainer, you need to do a practical translation."*

Affordances of Augmentation		
Subjectivity and prejudice of human judges:		
	Cognitive limitations of human brain	*"The human eye and human brain can't work so fast and accurate. There are too many decisions to be taken."*
	Human fatigue	*"When […] you're spending 14 hours a day in the gym, it's really hard to be fresh from the first moment of the first day until the last moment of the last day."* *"[T]he judges might be tired or thirsty or hungry, needing a break."*
	Human subjectivity	*"…we can make some mistake […] we have this subjectivity inside"*
Electronic systems' greater objectivity and impartiality		*"Judges can hear very often from the coaches that we've been biased with their athletes, and if the routine is evaluated by the system, who can you blame for low scores?"* *"We can make some mistakes about an objective thing, but the system can't."*
Subjectivity and prejudice of human judges:		
	Expectations of the judges	*"We anticipate, which helps you sometimes with your judgment."*
	Familiarity with the routine	*"If you're really familiar with the routine, it can influence your judgment positively or negatively."*
	Friendship between the judges and coaches	*"Sometimes judges and coaches [...] set up a good relationship."*
	The order of athletes' performance	*"[I]f you compete in the morning, judges are harder on you."*
	Unofficial guidelines	*"Judges have a certain average from a morning competition, and they need to keep this average between morning and evening scores."*
	Familiarity / expectations held for an athlete	*"[I]t's usually not fair when the judges know you and have seen you so many times during the training so they kind of know already where you will do a mistake."*
The higher accuracy of an AI-powered system:		*"What a human eye sees is one thing, but what the machine sees is more accurate."* *"The computer can do better, can better see angles, and it's more precise than human[s]."*
Lack of explanation		*"The judges don't tell us what exactly we get the deduction for."*
Potential increases in the judging process's explainability		*"When the system can provide some explanation or even a printed list of all deductions and scores […]– how the judgment was done, and everybody will understand everything."*
	Use for the inquiry	*"When they have something appealed [...] they will use this system to help them to evaluate, again, the whole routine."*
Speed-up judging process		*"The goal is to be able to help the judges […], and most importantly […] speed up the time of judging."*
Training to use the system		*The first step [...] is to open training camps where people test it and provide their feedback on it.*