

# The design of self-organizing human–swarm intelligence

Jonas D Hasbach<sup>1,2</sup> and Maren Bennewitz<sup>2</sup>

Adaptive Behavior

1–26

© The Author(s) 2021



Article reuse guidelines:

[sagepub.com/journals-permissions](https://sagepub.com/journals-permissions)

DOI: 10.1177/10597123211017550

[journals.sagepub.com/home/adb](https://journals.sagepub.com/home/adb)

## Abstract

Human–swarm interaction is a frontier in the realms of swarm robotics and human-factors engineering. However, no holistic theory has been explicitly formulated that can inform how humans and robot swarms should interact through an interface while considering real-world demands, the relative capabilities of the components, as well as the desired joint-system behaviours. In this article, we apply a holistic perspective that we refer to as joint human–swarm loops, that is, a cybernetic system made of human, swarm and interface. We argue that a solution for human–swarm interaction should make the joint human–swarm loop an intelligent system that balances between centralized and decentralized control. The swarm-amplified human is suggested as a possible design that combines perspectives from swarm robotics, human-factors engineering and theoretical neuroscience to produce such a joint human–swarm loop. Essentially, it states that the robot swarm should be integrated into the human's low-level nervous system function. This requires modelling both the robot swarm and the biological nervous system as self-organizing systems. We discuss multiple design implications that follow from the swarm-amplified human, including a computational experiment that shows how the robot swarm itself can be a self-organizing interface based on minimal computational logic.

## Keywords

Human–swarm interaction, intelligent system, bio-inspiration, swarm robotics, human-factors, theoretical neuroscience, self-organization, swarm-amplified human, distributed shortest path finding

Handling Editor: Farshad Arvin, The University of Manchester, UK

## 1. Introduction

*The division of the World's system into Natural and Man-made died with Darwin.*

Ross Ashby, in Conant (1981)

Swarm robotics is one of the grand frontiers in modern robotics (Yang et al., 2018). Robot swarms promise utility in multiple application domains such as search and rescue, military operations, pollution control and space exploration. In contrast to other multi-robot systems (Parker, 2007), robot swarms are designed to solve tasks by local robot interactions only, resulting in robustness against node failures and scalability of swarms (Balch & Parker, 2002; Barca & Sekercioglu, 2013; Bayındır, 2016; Brambilla, Ferrante, Birattari, & Dorigo, 2013; Dudek, Jenkin, Milios, & Wilkes, 1996; Hamann, 2018; Şahin, 2004; Schranz, Umlauf, Sende, & Elmenreich, 2020). Fascinatingly, simple interaction rules between agents can be sufficient to produce

complex collective behaviour (Bonabeau, Dorigo, Marco, Theraulaz, & Theraulaz, 1999; Camazine et al., 2003; Couzin, 2009). Swarm intelligence may, therefore, be implemented in simple robots as is the current standard in swarm robotics research. However, collective algorithms may also be useful in multi-robot systems consisting of capable robot agents.

One of the challenges to deploy robot swarms in the real world is the coupling of human operators with robot swarms. The need for human systems integration can be grouped into three categories. First, the swarm

<sup>1</sup>Human-Machine-Systems, Fraunhofer Institute for Communication, Information Processing and Ergonomics FKIE, Wachtberg, Germany

<sup>2</sup>Humanoid Robots Lab, Computer Science VI, University of Bonn, Bonn, Germany

### Corresponding author:

Jonas D Hasbach, Human-Machine-Systems, Fraunhofer Institute for Communication, Information Processing and Ergonomics FKIE, Fraunhoferstraße 20, 53343 Wachtberg, Germany.  
Email: [jonas.hasbach@fkie.fraunhofer.de](mailto:jonas.hasbach@fkie.fraunhofer.de)

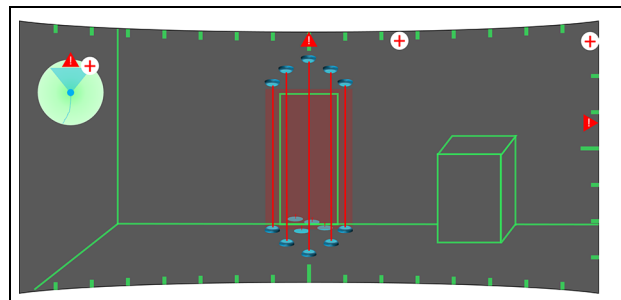
may not be capable of achieving its goals autonomously, resulting in the need to incorporate a human controller to compensate for the mismatch between reality and the robot's control model (Woods & Hollnagel, 2006). Second, increased autonomy may promote human-out-of-loop effects that may even result in reduced performance of the overall system (Bainbridge, 1983; Hollnagel & Woods, 2005; Woods & Hollnagel, 2006). Third, some applications require human supervision because of legal and ethical concerns (Verbruggen, 2019).

Human–swarm interaction (HSI; Kolling, Walker, Chakraborty, Sycara, & Lewis, 2015) as a research field combines human-factors engineering and swarm robotics to address the challenge how the distributed nature of robot swarms and the centralized control and feedback demands of humans can be combined into one loop. A human operator must be able to exert control over the swarm, which stands in opposition to the traditional concept of swarm intelligence that strives for solutions without leadership (Barca & Sekercioglu, 2013). In addition, flexibility and scalability of swarms pose high demands on operator capabilities (Kolling et al., 2015). In general, HSI can be subdivided into remote interaction, where operator and swarm are not located in the same environment, and proximal interaction, where operator and swarm are in a line-of-sight relationship (Goodrich & Schultz, 2008; Kolling et al., 2015). Most real-world applications that include proximal interaction, for example, search and rescue, will probably also be in need of a remote component for overall command and control.

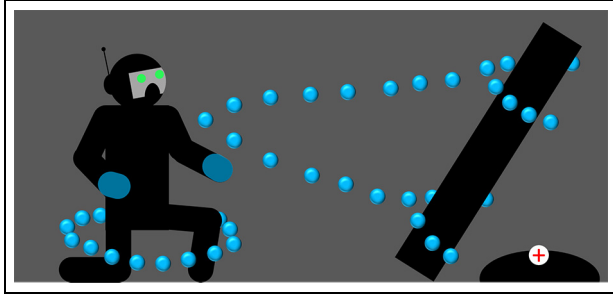
Around a decade ago, Penders et al. (2011) presented a prototype for HSI in the context of a fire-fighter application. In line with Penders et al., we often observe that humans are fully occupied with their tasks in real-world applications, such as in search and rescue, without the capacity to allocate mental workload to additional machine-supervision tasks. This has been formulated by Woods and Hollnagel (2006) as the law of stretched systems, which states that systems including humans tend to exploit their full capacity. Therefore, we envision a future human–swarm implementation where human operators are not required to deliberately control a robot swarm but where the swarm acts together with the human operator as one unity.

Imagine a future fire-fighter, who we shall call Tom, arriving at a warehouse emergency scene with a search and rescue assignment (Carrillo-Zapata et al., 2020; Payton, Estowski, & Howard, 2004; Penders et al., 2011). The warehouse is filled with smoke, resulting in poor visibility. Even worse, the deployed fire-responder drone could not be of much help because its sensors could not penetrate the specialized building material of the warehouse. Tom's order is to locate trapped humans, there may be four according to warehouse

security. Being in the warehouse, he must continuously assess the situation and balance trade-offs in this high-risk scenario. Fortunately, Tom is supported by a swarm of small robots that have already been deployed. Tom activates his capability amplifier which, besides communicating with human team members, amplifies his natural human sensory–motor range. The swarm robots spread into the whole warehouse (Payton, Daily, Estowski, Howard, & Lee, 2001; Penders et al., 2011), each individual being unaware of higher goals and solely driven by local stimulus–response rules (Arkin, 1998). After a short while, the sensor fusion related to life detection of one lucky robot activates, which triggers a recruiting message to other neighbouring robots that excitingly storm into the area and start exploring (Bashyal & Venayagamoorthy, 2008). Indeed, the sub-swarm concludes by collective estimation (Hamann, 2018; Valentini, Ferrante, & Dorigo, 2017) that it is likely for a human survivor to be present. The collective triggers a new message that is spread across the swarm ad hoc network (Payton et al., 2004) to Tom and the other fire-fighters. Tom reports to control that he is on his way and starts following his 'artificial sixth sense' that guides him into the direction of the possible victim (Payton et al., 2004; Penders et al., 2011). At the control terminal that is located in a fire truck standing outside the warehouse, the command and control operator Mary supervises possible paths for Tom as provided by the swarm estimation (Lee, Fekete, & McLurkin, 2016; Penders et al., 2011) while coordinating Tom's movement with other deployed forces. As Tom approaches a closed door, he suddenly feels being repelled by it, seemingly dragged back, induced by an increased stiffness of his body armour. His visual interface flags the door in a red threatening colour (Figure 1) and his whole front armour starts to vibrate. His head-up display shows that several of his cheap swarm companions have been destroyed only recently while going underneath the door, now forming a cemetery of melted robots. In a last act, the faithful robots propagated a distress signal (Hölldobler & Wilson, 1990) to other neighbouring robots that by their programmed instincts formed a



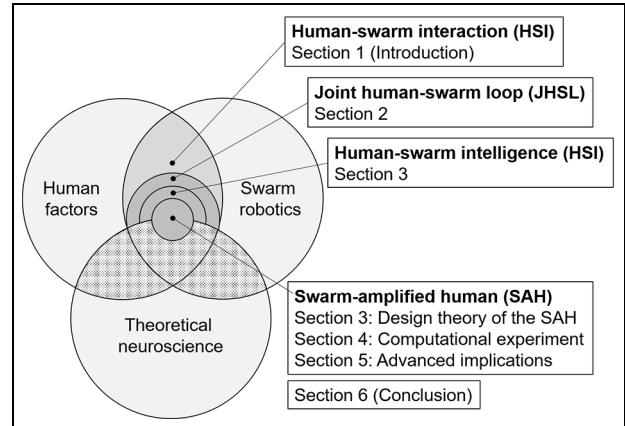
**Figure 1.** Sketch of Tom's head-up display showing the swarm forming a safety perimeter to warn Tom of a hazardous area.



**Figure 2.** Sketch of Tom being protected by the robot swarm while he manipulates an object with the swarm.

safety zone in front of the door. Tom starts following an alternative route to the possible victim. His heart-beat and respiration accelerate—even the advanced homeostatic temperature control of his body armour cannot make Tom’s physical and psychological stress go all away as he enters increasingly hazardous zones. With increasing stress (Schwarz & Fuchs, 2017), some robots that have been exploring the warehouse allocate to Tom (Villani, Capelli, Secchi, Fantuzzi, & Sabattini, 2020). It appears as if the robots sensed that Tom requires increased protection, as if they are part of Tom. The allocated robots form a safety mesh around him (Alboul, Saez-Pons, & Penders, 2008; Figure 2, left side), closely monitoring the local area and always ready to protect Tom from explosions or falling objects—even if the act would disable themselves. Finally, Tom reaches the survivor. To his distress, the victim is trapped under a heavy rack. Tom engages his hand-controlled swarm manipulator (Tsykunov et al., 2019); after a few seconds, different tiny robots start to form what looks like growing tentacles originating from Tom (Nouyan, Campo, & Dorigo, 2008) with handlike structures at their tips (Dorigo et al., 2013; Mathews, Christensen, O’Grady, Mondada, & Dorigo, 2017; Oh, Shirazi, Sun, & Jin, 2017; Rubenstein, Cornejo, & Nagpal, 2014; Slavkov et al., 2018), slowly lifting and holding the rack in a secure position (Figure 2, right side). In the meantime, Sarah, a fellow fire-fighter, has arrived and in close coordination both of them expertly provide first aid and transport the victim out of the building aided by the robot swarm (Dorigo et al., 2013; Gross & Dorigo, 2009). Outside, the patient is handed over to the medical staff on scene. Later that day, Tom is being told that the patient survived.

The above story highlights the ecological interdependence of task objectives, humans, robots, interfaces and environment in a real-world application. If we strive towards such a future exploitation of technology, it is essential that HSI combines the human and swarm perspectives in such a way that the overall system behaviour is tuned to the real-world tasks. This highlights the importance of a holistic-system view that can formulate how components of the human–swarm loop should be combined. The objective of the present work



**Figure 3.** Structure of this article and its embedment in the research landscape.

is to develop the theoretical foundations of such a systems view for HSI in terms of what we call a design theory, that is, a theory that formulates predictions about how human–swarm components should be combined so that desired joint-system dynamics emerge. In particular, we combine perspectives from swarm robotics, human-factors engineering and theoretical neuroscience to develop a design theory that utilizes the concept of self-organization. Emphasis is given to how work in swarm robotics relates to well-known concepts in neuroscience. We hope that this approach can serve as a path for integrating fragmented HSI solutions into one framework and may trigger the development of diverse HSI design theories.

The structure of this article and its embedment in the research landscape can be seen in Figure 3. Table 1 provides a glossary of central terms. In Section 2, joint human–swarm loops (JHSLs) are introduced as a holistic perspective on HSI. Section 3 defines the challenge for HSI as the design of human–swarm intelligence. In Section 4, the swarm-amplified human (SAH) is discussed as a possible design theory that builds on the concept of self-organization to join human and swarm. The feasibility of a self-organizing SAH architecture is demonstrated in Section 5 by means of a simple computational experiment. Section 6 provides advanced implications that follow from the SAH. Finally, Section 7 summarizes central insights and empathizes the benefit of developing design theories.

## 2. Joint human–swarm loop

### 2.1. Holistic perspective on HSI

JHSL refers to a cybernetic perspective on HSI where human and swarm are seen as interdependent parts of the same goal-oriented system (Ashby, 1961a; Clark & Chalmers, 1998; Hasbach, Witte, & Bennewitz, 2020; Hollnagel & Woods, 2005; Hutchins, 1995; Rasmussen, Pejtersen, & Goodstein, 1994; Woods & Hollnagel,

**Table 1.** Glossary of central terms (in order of appearance).

Term	Meaning
Robot swarm	Collective of relatively simple robots that solve tasks by local robot interactions only
Human–swarm interaction (HSI)	As a research field: Combination of human-factors engineering and swarm robotics that addresses the challenge how the distributed nature of robot swarms and the centralized control and feedback demands of humans can be combined into one loop As a system component: Interaction between human and robot swarm
Joint human–swarm loop (JHSL)	Cybernetic perspective on human–swarm interaction where human and swarm are seen as interdependent parts of the same goal-oriented system
Centralized–decentralized trade-off	Trade-off when implementing both centralized and decentralized control and communication in the same system
Design theory	A theory about what a good design would be
Robustness	In swarm robotics: The ability of a swarm to continue operation when disturbed, while particular focus is given to the loss of robot agents In human–swarm intelligence: A system that is capable of achieving its system purpose over a variety of situations given a particular internal organization
Flexibility	In swarm robotics: The property of a swarm to deal with different tasks In human–swarm intelligence: Underlying property of robustness, adaptation and scalability
Scalability	In swarm robotics: The capability to solve tasks over different swarm sizes In human–swarm intelligence: The capability to solve tasks over different swarm sizes and numbers of human operators
Adaptation	In swarm robotics: The term is not consistently used in the swarm robotics literature. It may refer to a change in behaviour, change by evolution, sensory tuning and learning In human–swarm intelligence: A change of the internal organization to continuously match the demands of the environment
Intelligent system	A system that chooses correct decisions given the present situation set
Situation set	Possible situations a system has to cope with
Internal organization	The interdependence of the system parts that provide the algorithms or input–output laws to cope with a set of situations (i.e. the agent's controller)
Human–swarm intelligence (HSI)	Interpretation of human–swarm interaction that refers to the learning joint human–swarm loop as a capable controller of environmental demands
Human–swarm intelligence challenge	The challenge of how to combine the centralized control demands of human operators with the decentralized control demands of swarms into one goal-seeking system while promoting sustained adaptive, robust and scalable behaviour
Central pattern generator (CPG)	Type of low-level neural network that produces rhythmic outputs for stereotypical locomotion while integrating bottom–up sensory feedback and top–down modulation signals
Abstract neural control architecture	Control architecture with semi-autonomous subsystems that are yet responsive to higher-order modulation signals
Swarm-amplified human (SAH)	As a design theory: Holistic design theory for human–swarm intelligence that treats the swarm as part of the human's low-level nervous system function As a system: Human that is extended by semi-autonomous robot swarms that are designed to function and feel like body parts
Subswarm	Behavioural subgroups $P$ of swarm robots that are defined by their respective locally perceived environment and global human modulation. Due to their semi-autonomy, they may also be referred to as local sensory–motor loops
Artificial body part	Subswarm (i.e. a behavioural subgroup $p$ of the robot swarm) that can be interpreted as an artificial body part by the human observer
Self-organization	The observable shift from unordered to ordered spatio-temporal patterns on the group level that are the result of local interactions on the individual level
Connectome graph $G_C$	Self-organizing connection network formed by the robot swarm that interfaces humans with environmental features based on $G_P$ .
Possible connections graph $G_P$	Connection network formed by the robots based on the current positions of the robots in space

2006). The focus is on how the components of the JHSL can be combined together so that the whole system displays desired behaviours in given situations. We choose the term JHSL to stress that HSI as a research discipline has to consider the whole human–swarm loop.

In its simplest form, a JHSL is conceptualized as made of three interdependent components (i.e. human, swarm and interface) placed in a certain environment, while each component relates to a given perspective for design (Figure 4). From the human perspective, which may be called human-factors engineering or human-centred design, one must investigate the following questions:

- What human needs must be satisfied to allow them to achieve their goals efficiently when part of the socio-technical system (Carrillo-Zapata et al., 2020; Chen & Barnes, 2014; Lewis, Sycara, & Walker, 2018; Nam, Walker, Li, Lewis, & Sycara, 2019; Penders et al., 2011; Podevijn et al., 2016; Roundtree, Goodrich, & Adams, 2019)?
- How do the operators tend to use the swarm to align it with their goals (Walker et al., 2012; Woods & Hollnagel, 2006)?
- How do operators coordinate goals and actions with other agents of the socio-technical system (Coucke, Heinrich, Cleeremans, & Dorigo, 2020; Ellwart, 2011; Ellwart, Happ, Gurtner, & Rack, 2015; Klein, Woods, Bradshaw, Hoffman, & Feltovich, 2004; Tavakoli, Nalbandian, & Ayanian, 2016; Woods & Hollnagel, 2006)?
- How does the human operator's mental model about swarm control tend to change over time, and how should the operator be trained in swarm control (Hasbach et al., 2020)?

From the swarm perspective, which may be called swarm engineering, one must investigate the questions:

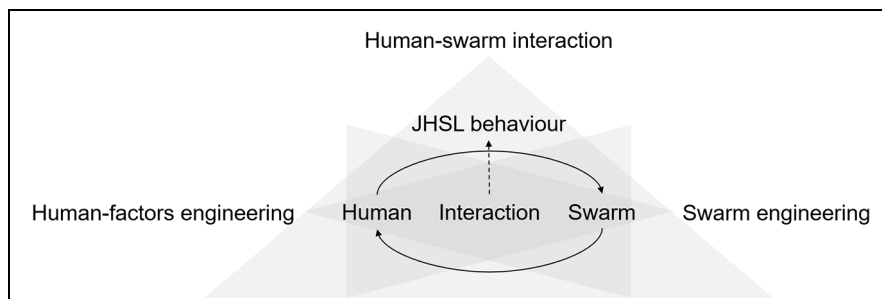
- What socio-technical requirements must be satisfied so that given swarm implementations can achieve their goals efficiently (Bashyal & Venayagamoorthy, 2008; Brambilla et al., 2013;

Hamann, 2018; Kolling et al., 2015; Schranz et al., 2020)?

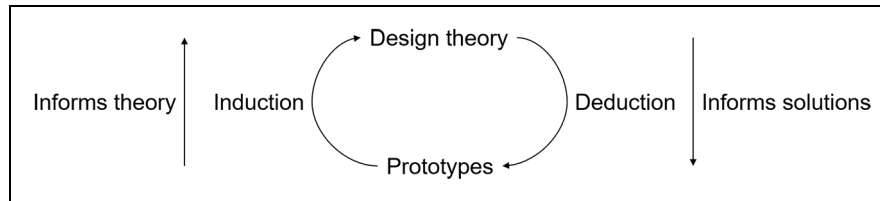
- How can the swarm robots coordinate goals and actions with other agents of the socio-technical system (Coucke et al., 2020; Hamann, 2018)?
- How can the robots' control models be updated during development (Birattari et al., 2019; Brambilla et al., 2013; Hamann, 2018; Nolfi & Floreano, 2000) and deployment?

Finally, HSI joins human and swarm into one goal-oriented system. The design objective is to find an interaction solution that combines the strengths of humans and swarms so that the joint-system capabilities that cope with the application scenario are maximized while reducing inner complexities that do not contribute to goal achievement. Thus, the JHSL behaviour is the desired level of performance analysis (Hollnagel & Woods, 2005). A failure to produce a joint-system relates to a violation of the law of requisite variety (Ashby, 1961a); a disjointed system produces variety in the internal human–swarm loops that is in conflict with producing variety that copes with the task objectives the system was actually built for (Hollnagel & Woods, 2005). Put simply, the swarm should not become unreasonably complex for the human operator, as well as the other way around, wasting resources that are needed to deal with actual tasks. From the HSI perspective, one must investigate the additional questions:

- What is the overall JHSL objective and what is required to solve it efficiently (Carrillo-Zapata et al., 2020; Cooke, Gorman, Myers, & Duran, 2013; Hollnagel & Woods, 2005; Hutchins, 1995; Penders et al., 2011; Rasmussen et al., 1994)?
- How can the capabilities of humans and swarm be combined to solve the JHSL task including task sharing and task allocation (Brown, Jung, & Goodrich, 2014; Coppin & Legras, 2011; Hutchins, 1995; Kolling et al., 2015; Musić & Hirche, 2017; Nam et al., 2019; Parasuraman, Sheridan, & Wickens, 2000; Walker et al., 2013)?
- How can the JHSL coordinate with external systems (Rasmussen et al., 1994)?



**Figure 4.** A human–swarm loop can be approached from different perspectives. HSI should consider the whole JHSL.



**Figure 5.** Design as part of the empirical cycle requires the development of a theory about what would be a good design. The theory should be subjected to empirical probing by evaluating prototypes that implement the theory.

- How can the communication link between human and swarm be established (Balch & Parker, 2002; Barca & Sekercioglu, 2013; Kolling et al., 2015)?
- How can humans issue control commands (Kolling et al., 2015)?
- How can swarms be controlled by external inputs (Kolling et al., 2015)?
- How can information about the environment and the swarm be fused to produce global information (Hamann, 2018; Kolling et al., 2015)?
- How can this information be presented to humans (Kolling et al., 2015)?

## 2.2. Centralized–decentralized trade-off

Each facet of the JHSL must be designed in relation to the joint-system goals as well as the capabilities and demands of the other components. Probably most importantly, a centralized control and communication approach may be a reasonable design solution from the human’s perspective but ignores the benefit of distributed control and communication provided by the swarm intelligence approach (Barca & Sekercioglu, 2013; Crandall et al., 2017). After all, integrating a human operator in a fully distributed system makes him a potential single point of failure. On the contrary, there is a need for the human to be part of the system loop after all. Therefore, balancing between centralized and distributed computation may be defined as the main challenge for combining human and swarm into one efficient joint-system. From a human-factors engineering perspective, this challenge can be seen as a special case of human–autonomy teaming via designing (flexible) levels of automation (Brown et al., 2014; Chen & Barnes, 2014; Coppin & Legras, 2011; Musić & Hirche, 2017; Nam et al., 2019; Parasuraman et al., 2000; Walker et al., 2013). For example, a shared control between human operator and swarm has been proposed to dynamically moderate the impact of human agents on the swarm (Ashcraft, Goodrich, & Crandall, 2019; Crandall et al., 2017). The challenge also relates to investigations on hierarchical control in swarm robotics, where a central robot with superior information is part of the control architecture (Dorigo, Theraulaz, & Trianni, 2020; Mathews et al., 2017; Zhu et al., 2020). From the literature, an underlying theme emerges where both a central unit with broad

capabilities (i.e. humans) and decentralized units with narrow capabilities (i.e. swarm robots) have, sometimes partly, a stake in decision-making processes to balance the centralized–decentralized control trade-off.

## 3. Human–swarm intelligence

### 3.1. Precondition for a design theory

Given the interdependence of design solutions for human, robot swarm and interface, in the following a holistic approach on how to best join human and swarm into one goal-directed loop is developed. Such a design theory as part of the empirical cycle must formulate testable predictions about what combination of design solutions are beneficial (Figure 5). While classical scientific theories model natural phenomena, a design theory models what would be a good design. In this framework, design prototypes are seen as hypotheses deduced of the underlying design theory (Woods & Hollnagel, 2006).

A precondition for the construction of a design theory is a definition of desired JHSL behaviour. While most joint-system goals depend on the application scenario, there are three system properties of HSI that seem desirable independent of the application; adaptation, robustness and scalability. We selected these variables guided by Beer’s aphorism: ‘*The purpose of a system is what it does*’ (Beer, 1985, p. 99) or rather in the context of design ‘...what it is intended to do’. Adaptation, flexibility and robustness are important concepts in human–machine loops (Hollnagel & Woods, 2005; Woods, 2015, 2018; Woods & Hollnagel, 2006), where, in particular, flexibility and robustness are also of interest in swarm robotics (Şahin, 2004). In addition, robot swarms are designed to be scalable (Şahin, 2004).

There are no globally accepted definitions of robustness, flexibility, scalability and adaptivity across different scientific fields. According to the seminal work of Şahin, (2004), the following definitions can be given from the perspective of swarm engineering. Robustness is the ability of a swarm to continue operation when disturbed, while particular focus is given to the loss of robot agents. Flexibility is the property of a swarm to deal with different tasks. Scalability is the capability to solve tasks over different swarm sizes. The term

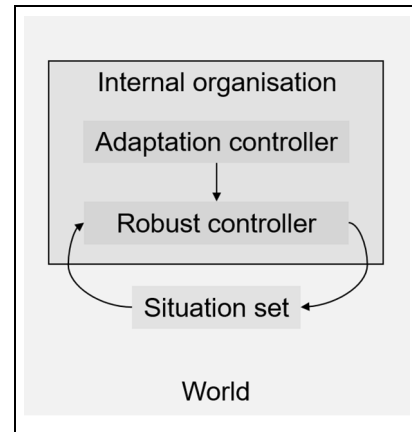


‘adaptation’ is not used by Şahin and seems to have different meanings in the swarm robotics literature. In behaviour-based robotics, which is closely related to swarm robotics, adaptation may refer to behavioural changes, evolution, sensory tuning and learning (Arkin, 1998). As we investigate the JHSL rather than the swarm as an isolated system, we will reassess the desired properties in the following to fit the joint-systems view.

### 3.2. HSI as human–swarm intelligence

In general, one wants any intelligent system to choose the correct decision given the present situation set (Ashby, 1960, 1961b). Let a system be called robust if it is capable of achieving its system purpose, that is, choosing correct decisions out of a set of possible decisions, over a variety of situations given a particular internal organization (Ashby, 1960; Woods, 2015). The internal organization is defined as the interdependence of the system parts (Ashby, 1947; Conant, 1976) that provide the algorithms or input–output laws to cope with a set of situations. It may also be referred to as the agent’s controller. By applying Beer’s aphorism (Section 3.1), it follows that the internal organization encodes the system’s goals in terms of possible attractor states, or rather the corresponding selected parameter set, of the system (Ashby, 1960). In terms of stability, a robust system can absorb particular shocks from the environment (Woods, 2015) without driving essential variables out of their tolerance values (Ashby, 1960). However, a particular internal organization can only deal with a limited set of situations when faced with complex or dynamic environmental demands (i.e. when environmental variety is high; Ashby, 1961a; Woods, 2018). To keep demands in check that are not modelled by the robust system, while modelling is required as stated by the good regulator theorem (Conant & Ashby, 1970; Seth, 2014), the system must be adaptive, that is, it must be able to change its internal organization to contentiously match the demands of the environment (Ashby, 1960; Friston, 2010; Nikolić, 2015; Powers, 1973; Russell & Norvig, 2016; Woods, 2018). If a system can continuously adjust its internal organization to fit new opportunities and demands, it is referred to as resilient (Woods, 2015).

Figure 6 shows the desired intelligent system as two abstract controllers (Ashby, 1960; Seth, 2014) placed in the world. By abstract, we mean that the two loops may not be observed as distinct matters and functions in the real system. The first-order robust controller can autonomously cope with the inputs of the world to produce desired internal and external states. In general, by connecting the robust controller to the world, the system behaviour ‘flows’ as defined by the robust controller logic and environment (Nikolić, 2015). However, the robust controller alone will ultimately fall short



**Figure 6.** An intelligent system is abstracted as an internal organization that includes two types of controllers. The robust controller has the capability to cope with possible situations while the adaptation controller improves the robust controller’s capabilities.

when faced with the variety of the world. Therefore, the second-order adaptation controller implements rules that adjusts the robust controller when the fit between the robust controller and the situation set is not optimal. While the robust controller may switch between encoded system goals (i.e. switch between attractors in the state space as given by the chosen robust controller and environment), the adaptation controller may also generate new system goals by changing the robust controller (i.e. change the robust controller’s state space). The adaptation controller may be said to be robust itself if it correctly adjusts the robust controller over a variety of situations.

By imagining a kitten approaching a fire, Ross Ashby (1960) has provided us with an illustrative example for intelligent behaviour. If a kitten gets too close to a fire for the first time, the unpleasant stimuli may trigger a reorganization of the internal neural machinery leading to a behavioural avoidance of this situation in the future. Thus, the kitten’s adaptation controller, here in the form of operant conditioning, has changed the robust controller that deals with the behavioural response to the situation ‘fire’ from ‘approaching fire’ into ‘avoiding fire’. After the kitten has learned to avoid fire, an observer may say the kitten is robust because it possesses the means to cope with the situation. Note how robustness does not require an update in the kitten’s internal organization; it is a property that the kitten already has at its disposal. Importantly, both adaptation and robustness ensure stability of essential variables, that is, the survival of the kitten, and both do so over a range of situations. What differs is the underlying mechanisms: the robust controller deals with expected (modelled) surprises, while the adaptation controller deals with unexpected (unmodelled) surprises (Ioannou & Sun, 1995; Rasmussen et al., 1994; Woods, 2018). Of course, in biological systems, we may

not be able to differentiate clearly between the adaptation controller and robust controller as well as modelled and unmodelled surprises.

In HSI, a JHSL learning that swarm agents have a high probability to be destroyed in a given area and, therefore, updates its internal organization to cope with the situation implements true adaptivity. The update of the internal organization can either refer to the human's control model of the world (Friston, 2010; Hollnagel & Woods, 2005; Johnson-Laird, 1983), or to the robots' control models (Arkin, 1998). In the former, the human changes his control behaviour while making use of the robot swarm's capabilities, for example, by manually steering the swarm away from the area. In the latter, the robots' capabilities are updated, which enable the robot swarm to autonomously avoid the area. It should be noted that both forms of adaptation lead to the desired change in JHSL behaviour. When the JHSL is observed at a later time, we may call it robust because it is behaviourally avoiding the loss of robots while being tolerant to certain changes in the environment. If robots are lost in that area and the JHSL can accomplish its goals nonetheless, the JHSL displays another strategy leading to robustness that corresponds to the view of robustness in swarm robotics.

Some situations that the JHSL may have to cope with are highly demanding, resulting in the need that goals of the systems need to be balanced. For example, the JHSL could be faced with an unmodelled exploration–exploitation trade-off. An intelligent JHSL will not keep this trade-off constant but apply a search strategy to find an acceptable exploration–exploitation balance. If observed, the JHSL may look like switching between different behavioural states in its attempt to narrow down plateaus in the solution landscape. It is said that the adaptation controller searches for an acceptable solution for the robust controller. The adaptation controller may even ignore one goal for a time period by, for example, arguing that attempting to reach both goals simultaneously will result in both goals being out of reach. Note how adaptation can relate to cognitive reasoning (Hollnagel & Woods, 2005).

Finally, the third desired system property is scalability that can be seen as a special case of robustness as defined above. In swarm robotics, it is the capability to adjust the number of robots during deployment, such as two separated bird flocks forming interactions to become one flock. In HSI, swarm scalability is challenging because a human operator has a finite capacity to supervise independent robots (Kolling et al., 2015; Lewis, 2013). Scalability can also refer to the possibility to control the swarm with a different number of human operators working as a team (Coucke et al., 2020; Grosh & Goodrich, 2020). This makes the combination of proximal and remote interaction a special case of scalability, as it essentially combines human operators

with different ranges of skills, situational awareness and control possibilities.

From the discussed point of view, flexibility underlies all three discussed properties (robustness, adaptation and scalability) and is, therefore, not explicitly formulated as a desired joint-system behaviour. A robust controller is flexible in the sense that it applies different solutions to deal with the situation set. An adaptation controller results in system flexibility by increasing the range of the robust controller. In addition, the robust controller itself must be flexible so it can be adjusted by adaptation. Finally, scalability as a special form of robustness results in the flexibility over different swarm size situations.

To conclude, HSI may also be read as human–swarm intelligence referring to learning JHSL as a capable controller of environmental demands. Note that the term also relates to human swarm intelligence as human collectives that display swarming behaviour (Sumpter, 2006), thereby stressing the importance of integrating the human into the swarm (Kolling et al., 2015) as well as human teamwork (Coucke et al., 2020; Tavakoli et al., 2016). Now the overall challenge of HSI can be defined as combining the centralized control demands of human operators with the decentralized control demands of swarms into one goal-seeking system while promoting sustained adaptive, robust and scalable behaviour. A candidate design theory of JHSL must provide a solution to this challenge, that is, formulate predictions about how the desired properties can be designed.

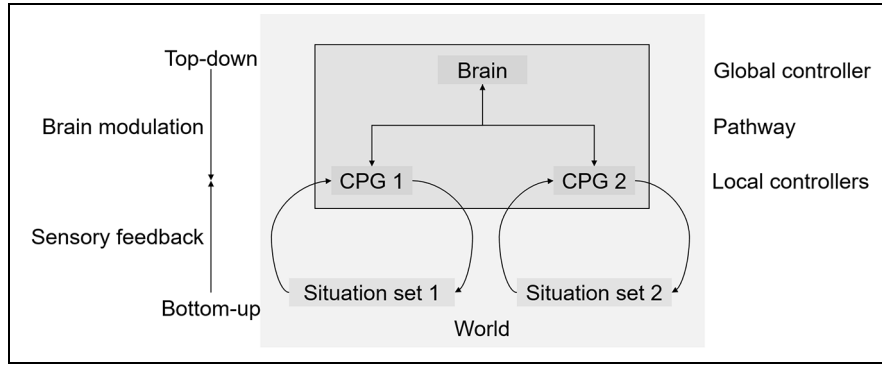
## 4. Design theory of the SAH

### 4.1. Neural control architecture

The main challenge was formulated as how centralized and decentralized control can be brought together while preserving the benefits of both. As humans are biological systems and swarm robots are heavily bio-inspired, we searched for bio-inspiration that can inform how to cope with the centralized–decentralized control trade-off. We found such bio-inspiration in the abstract neural control architecture. This is a good start, because the nervous system is assumed to be a capable controller of environmental demands, that is, it implements the good regulator theorem (Friston, 2010; Seth, 2014).

In biological nervous systems, stereotypical locomotion is generated by a special form of neural networks that are located at the low-level, or at the 'front end', of the nervous system. These neural networks are called central pattern generators (CPGs) (Arbib, 1989; Ijspeert, 2008; Marder & Bucher, 2001; Pearson & Gordon, 2013). CPGs are examples of semi-autonomous distributed controllers that are robust to a set of situations. While the correct CPG function depends on sensory feedback, higher-order ('back end')





**Figure 7.** Central pattern generators (CPGs) illustrate the abstract neural control architecture. CPGs are low-level neural networks that produce rhythmic output for stereotypical locomotion while making use of sensory feedback. At the same time, CPGs are sensitive to modulation signals from higher-order systems. This architecture balances between centralized and decentralized control.

vertically descending brain signals may also modulate the activity of the lower-order CPGs (Figure 7). Thus, at least some neural circuits of the biological sensory–motor loop seem to be capable of producing robust stereotypical sensory–motor behaviour while sensitive to higher-order neural modulatory signals. Imagine walking down the street. When walking, you rarely think about walking. Instead, the stereotypical locomotion is taken care of by your low-level sensory–motor loop. Yet, you can easily choose to shift from walking to running when you desire to do so; your higher-order system modulates the operation of the hidden low-level sensory–motor loop.

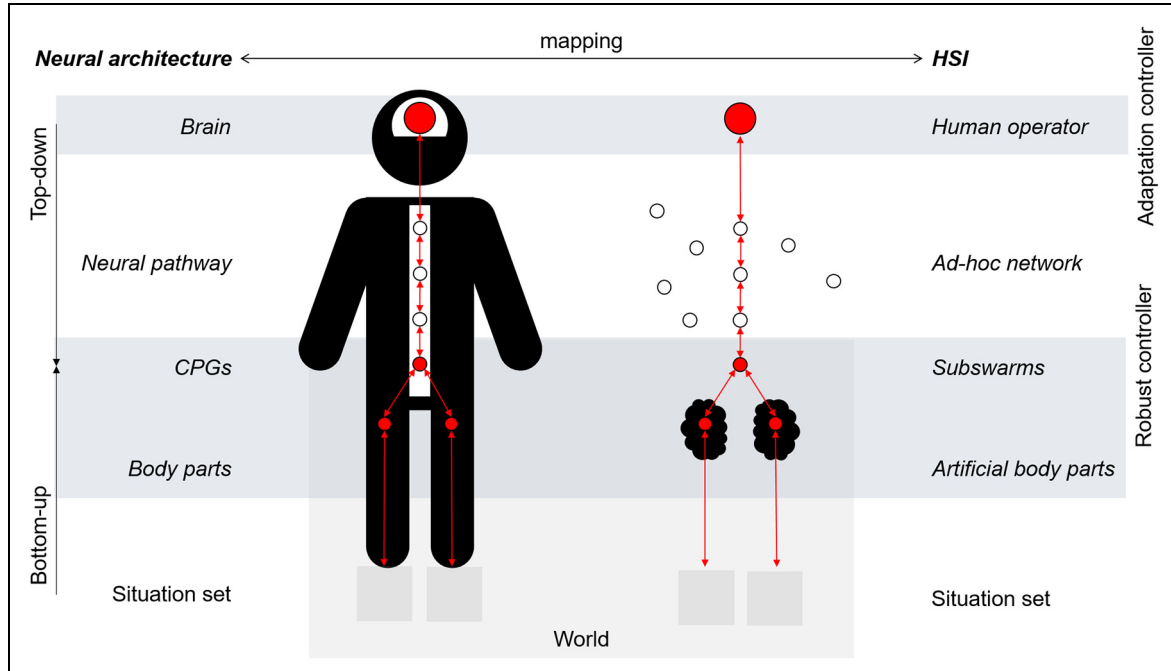
This control architecture with semi-autonomous subsystems that are yet responsive to higher-order modulation signals may be a fundamental principle in the control of artificial and natural complexity. To the roboticist, it is known as the hybrid deliberative/reactive architecture (Arkin, 1998; Murphy, 2019). It is a principle known in human–robot interaction as well. For example, a space telerobot must implement sufficient semi-autonomy while at the same time responsive to delayed inputs from human operators on earth (Sheridan, 1992). The creation of semi-autonomous subsystems is also relevant in the steering of human organizations (Senge, 2006).

#### 4.2. Neural architecture applied to HSI

We think this abstract neural control architecture may be also mapped onto the design challenge of HSI (Figure 8), because a higher-order controller (*brain*  $\mapsto$  *human operator*) modulates a distributed and semi-autonomous system (*CPG*  $\mapsto$  *subswarm*) where the signals are transmitted over a pathway (*neural pathway*  $\mapsto$  *ad hoc network*). Therefore, we propose to treat the swarm as low-level sensory–motor loops, as artificial swarms are better seen as task-specific robust controllers when compared to the broad capabilities possessed

by humans (Brooks, 1990; Woods & Hollnagel, 2006), while human operators are seen as adaptation (i.e. cognitive, Hollnagel & Woods, 2005) controllers.

The main advantages of this control architecture can be observed in biological systems. CPGs represent a fast controller that will, however, fail when faced with the complexities of the world, while brain cognition<sup>1</sup> represents a slow controller that is much more capable of dealing with world complexities. Imagine again walking down the street. A slow feedback controller may not be able to efficiently stabilize your locomotion. The CPGs, on the contrary, which are located in proximity to the body’s sensors and effectors, can react much faster. The CPGs are ‘outsourced’ from the brain’s perspective to establish fast (short) feedback loops, which decreases time synchronous information bandwidth needs between the brain and the body’s local sensor–effector systems (Ijspeert, 2008). Importantly, the brain as a controller of body states is confronted with less internal variety it needs to cope with, that is, the architecture minimizes internal control complexity. However, when you continue walking down the street you see an ice-cream truck that induces the response of excitingly running towards the potential food source. Classifying the ice-cream truck, adjusting your goal and updating the locomotion behaviour requires much more sensory integration and processing than producing stable stereotypical locomotion and, therefore, requires more resources and time. In addition, advanced anticipatory control may be needed to plan how to accomplish your goal of finally getting to enjoy your ice-cream. Importantly, it is the interplay of the fast but heuristic and the slow and sophisticated system that makes you a capable controller to deal with the world (Kahneman, 2011). Optimally, the slow-sophisticated controller implements learning mechanisms, that is, it is also an adaptation controller. Similarly, the swarm is part of the heuristic controller that can react fast to local changes (similar to walking on an uneven pavement)



**Figure 8.** Mapping from the abstract neural control architecture to HSI gives the SAH. In the SAH, the swarm is treated as a neural extension of the human's low-level sensory-motor loop. The swarm organizes into semi-autonomous subswarms that are sensitive to modulation signals from the human operator. For the human, these subswarms take the form of artificial body parts.

but requires cognitive modulation for complex and unmodelled scenarios. By implementing short feedback loops between subswarms and local environmental features, the required coordination between human operator and subswarm is reduced.

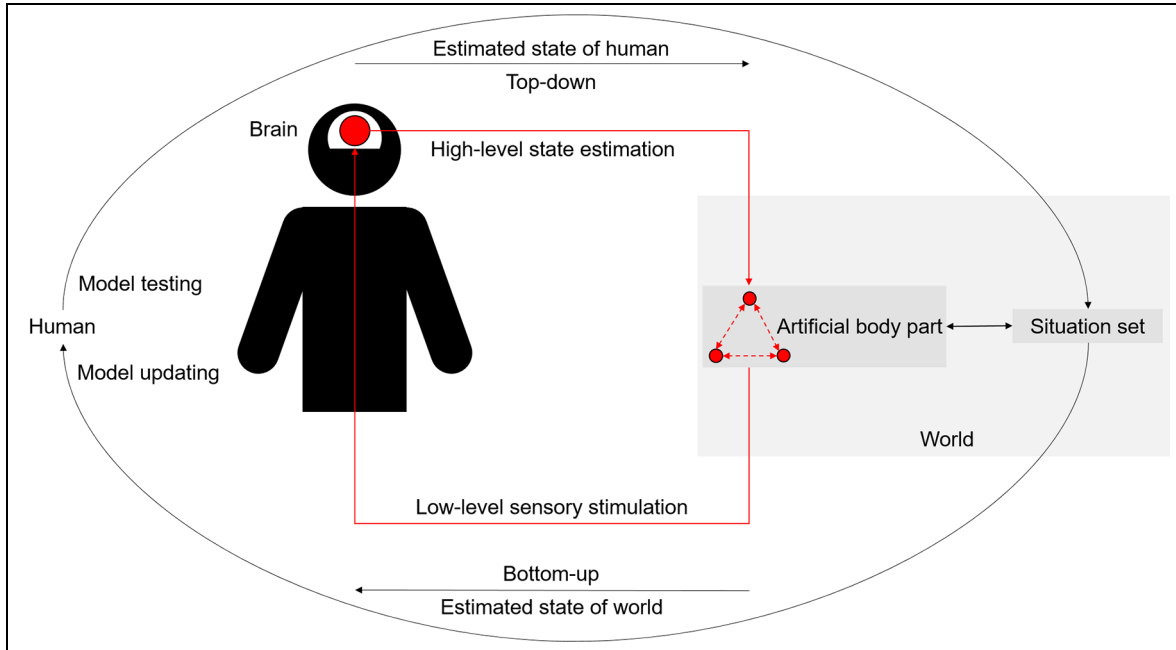
Based on the defined mapping *neural architecture*  $\mapsto$  *HSI*, the design theory that we would like to call the swarm amplified human treats the swarm as an extension of the human nervous system, integrated at low-level sensory-motor behaviour (Figure 8, right). From the human's perspective, the swarm then becomes an extension of the human body (Gibson, 1979; Hollnagel & Woods, 2005; Tsykunov et al., 2019); an artificial body part (*body parts*  $\mapsto$  *artificial body parts*). This shifts the focus from human-swarm channels to the interaction between human and environment, which is interfaced by the transparent swarm (Hollnagel & Woods, 2005; Le Goc et al., 2016; Sheridan, 1992). The swarm is, therefore, seen as a potential part of the human's perception-action cycle (Clark & Chalmers, 1998; Friston, 2010; Hollnagel & Woods, 2005; Neisser, 1976). From a swarm's perspective, robot swarms form behavioural subgroups ('artificial body parts') working in parallel that are, similar to CPGs, influenced by their respective locally perceived environment and global human modulation. The latter is conveyed by the swarm acting as a distributed neural pathway overlay (i.e. a flexible ad hoc network).

This swarm perspective can be viewed as an application of the mergeable nervous system concept in swarm robotics (Mathews et al., 2017) where a central unit,

here the human operator instead of a robot unit, exerts high-level control over its 'mergeable swarm body'. The SAH defines the vital extension that the subswarms can directly react to stimuli of the environment without the need that the signals are processed first at the central unit. One may say that the SAH includes inborn reactive pathways (i.e. stimulus-response mappings) in addition to modulation by higher-order units. From an evolutionary point of view the higher-order neural correlates (i.e. the neocortex) are younger than the lower-order neural correlates (e.g. reflex pathways). This highlights how higher-order units build on lower-order capabilities to make a capable controller of environmental complexities (Braitenberg, 1986; Brooks, 1986; Dawkins, 1976; Minsky, 1988; Nikolić, 2015; Pfeifer & Bongard, 2006; Powers, 1973).

#### 4.3. SAH as human-swarm intelligence

Human-swarm intelligence was defined in Section 3.2. In the SAH, the swarm's distributiveness is joined with the central control of the human after the blueprint of CPG control in the biological nervous system. Similar to CPGs, the swarm is capable of reacting quickly to the environment in a closed-loop (i.e. cope with some situations autonomously), while being biased by the human component. The swarm therefore is influenced by the human without fully being centralized to it. This property is demonstrated by taking the human out of the loop: the subparts of the swarm will, though with degraded performance, still show purposeful



**Figure 9.** In the SAH, the swarm is part of the human's perception–action cycle, which supports the update of the human's control model about the world (outer loop). Treating the swarm as part of the body relates to utilizing high-level human state estimations as swarm modulation signals and to providing sensory feedback consistent with the feeling of body stimulation.

behaviour. A morbid but known example is the headless chicken running around, although neuroscientific investigations in that area focus more on other animals like cats (Pearson & Gordon, 2013).

The SAH architecture is expected to result in adaptive, robust and scalable JHSL behaviour. Sustained adaptivity is the consequence of equipping the human operator with native high-level control over his swarm 'body', as the human component is the most capable adaptive system known to us. Thus, the biological brain is supported as joint-system adaptation controller by design, a solution that is applied by cognitive systems engineering (Hollnagel & Woods, 2005; Rasmussen et al., 1994; Vicente & Rasmussen, 1992; Woods & Hollnagel, 2006). Integrating the swarm at the low-level nervous system amplifies cognition as the latter builds on lower-order capabilities (Section 4.2). In other words, the cognitive system is provided with more task-relevant variety it can use for model update and goal-achievement (Ashby, 1961a; Friston, 2010); see outer loop in Figure 9. Given that neural systems feature high adaptive capability, implementing the swarm based on neural principles aims at further strengthening system adaptivity. Robustness and scalability is created by treating the swarm as a distributed neural machinery that excursively relies on local interactions. In neural terms, robustness comes in the form of plasticity (i.e. flexibility by adaptation) and graceful degradation to neuron failures. Neural scalability is vividly observed when comparing nervous systems across species.

#### 4.4. The human operator's perspective

From the human's perspective, treating the swarm as part of its body could provide a feasible design solution for controlling a complex system without the need for the swarm to explain itself. Humans are, assuming robustness of the neural implementation, not aware of the complexity of their own low-level neural machinery (Doyle & Csete, 2011; Kahneman, 2011). Yet it feels controllable and often effortless to us to produce complex behavioural patterns like riding a bike, dancing or making music. Therefore, functionally integrating the swarm into the nervous system may allow operators to develop an intuition of swarm control similar to their own body control.

Interfacing at the low-level nervous system means essentially two things (Figure 9). First, as said before, low-level stereotypic sensory–motor control often feels automatic to the conscious mind. When walking down the street, you rarely think about walking. Therefore, the (cognitive) state of humans should be translated by the interface into commands (Alboul et al., 2008; Bales & Kong, 2017; Karavas, Larsson, & Artemiadis, 2017; L Mondada, Karim, & F Mondada, 2016; Podevijn et al., 2016; Pourmehri, Monajjemi, Vaughan, & Mori, 2013; Schwarz & Fuchs, 2017; Villani et al., 2020) for the swarm clusters. This is referred to as passive (Kolling et al., 2015) or implicit (Gildert, Millard, Pomfret, & Timmis, 2018) interaction, that is, unconscious control, which can be contrasted to active/explicit control (e.g. deliberate gesture control;

Alonso-Mora, Lohaus, Leemann, Siegwart, & Beardsley, 2015; Gromov, Gambardella, & Di Caro, 2016; Nagi, Ngo, Gambardella, & Di Caro, 2015). The swarm forming a protective layer based on Tom's distress is an example of passive interaction (Section 1). A further design restriction is that the human operator perceives the passive translation as deterministic (Braun et al., 2018; Woods & Hollnagel, 2006). After all, you can deliberately control leg movement if you desire to do so. This feeling of being in control over the swarm clusters (alias artificial body parts) relates to human-factor issues like trust in automation (Lewis et al., 2018; Nam et al., 2019). Note that passive interaction may be used deliberately by a human if he knows the mapping between his own states and swarm states. The interface may additionally implement deliberative control of the artificial body parts, in line with how body control feels (Tsykunov et al., 2019), that can manually be triggered by the human operator. One may think of the swarm forming handlike structures where the individual swarm-fingers can be controlled by Tom.

Second, feedback should be provided via the bottom-up route similar to how sensor stimulation of the body feels. Take again the previous example of approaching a fire. While blindfolded, you are able to locate the source of the fire by moving around and observing the gradient of the perceived heat (Bashyal & Venayagamoorthy, 2008; Braitenberg, 1986; Lilienthal & Duckett, 2004; Nouyan et al., 2008; Schmickl & Crailsheim, 2008). The example illustrates the active nature of exploration in the perception-action cycle where sensory inputs regulate action outputs that, in turn, alter sensory inputs (Neisser, 1976). A similar loop may be implemented by a wearable interface that adjusts its feedback based on the estimated stimulus strength and location (Penders et al., 2011). The cognitive capability, that is, updating a model of fire location, is amplified by providing an increased sensory-motor range at the low level. Importantly, this fits to the swarm robotics' behaviour-based approach where stimulus class and strength is used to adjust behaviour (Arkin, 1998). Note also how this example relates to the real-world scenario of fire-fighter Tom.

By closing the loop as shown in Figure 9, the swarm potentially becomes part of the human's perception-action cycle. The SAH, therefore, relates to the notion of a cyborg that '[...] *deliberately incorporates exogenous components extending the self-regulatory control function of the organism in order to adapt it to new environments*' (Clynes & Kline, 1960, p. 27). To summarize, treating swarms as part of the human's low-level nervous system and modulation of semi-autonomous sub-swarms that form the interface to the environment are the very essence of the SAH to produce a JHSL.

#### 4.5. Self-organization as the glue between humans and swarm

Based on the engineering goal as put forward by the SAH to functionally integrate the swarm at the low-level nervous system, a concept is needed that allows to translate computational principles between nervous system and artificial swarm. Self-organization is a concept well-known to both neuroscience and swarm intelligence and has been proposed as an underlying computational mechanism for both systems (Couzin, 2009; Trianni, Tuci, Passino, & Marshall, 2011). It is defined as the observable shift from unordered to ordered spatio-temporal patterns on the group level that are the result of local interactions on the individual level (Ashby, 1947, 1962; Bonabeau et al., 1999; Camazine et al., 2003; Couzin, 2009; Hamann, 2018; Sumpter, 2006; Trianni et al., 2011). The individual level and their interactions refer to neurons as well as swarm agents, for example, ants, bees or termites, that form global networks based on local rules. Capabilities at the group level are observed that cannot be explained by neurons or ants in isolation. This system property is often paraphrased as 'more than the sum of its parts', stressing the importance of interactions between components.

Intriguingly, neural networks and biological swarms seem to resemble not only on this abstract level. On the microlevel, for example, both individual neurons and ants may have a refractory period, making activation unlikely shortly after being active (Couzin, 2009). On the macrolevel, swarms may follow psychophysical laws previously thought to be an exclusive property of cognitive systems (Passino, Seeley, & Visscher, 2008; Reina, Bose, Trianni, & Marshall, 2018; Sasaki & Pratt, 2018). It is an attractive hypothesis that neural and swarm decision-making share the same computational principles that have been manifested by evolutionary selection in different connected systems to cope with environmental complexity.

In the following, we will use self-organization in swarms and neural networks as common computational ground, as the glue, to join human and swarm into one goal-directed system. Rather than claiming that neural networks and biological swarms share the actual same computational basis, we argue that both can be modelled as self-organizing systems. Thus, both humans and swarms are abstracted into one common computational reference frame (Marr, 1982) to integrate the swarm at the low-level of neural sensory-motor loops. On this abstracted level, one may think of the human operator himself as a distributed self-organizing neural system that is extended by a neural-swarm machine. The benefit for swarm engineering by transferring neural principles to swarm networks that goes beyond the usage of neural robot controllers has been demonstrated recently by the Bruessel's group (Mathews et al., 2017; Zhu et al., 2020), Monaco, Hwang, Schultz, & Zhang (2020), and Otte (2018).

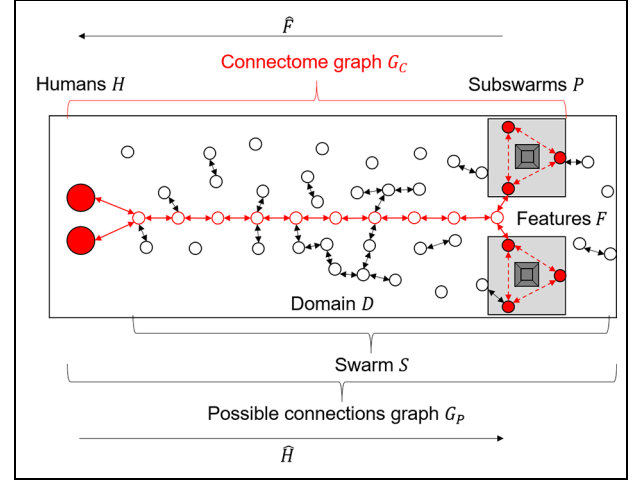
#### 4.6. SAH as self-organizing HSI

When the swarm is the interface between the variety of human cognition and the world, it becomes a potential bottleneck. The swarm must, therefore, be designed to allow for the sufficient and reliable conductance of information between human and situation set (Ashby, 1961a), while the required bandwidth for signal forwarding is minimized by implementing local sensory–motor loops. A possible solution lies in the application of self-organizing principles for pathway formation that, based on local interactions, allows for scalable parallel processing.

The SAH loop  $L = \{H, S, G_C, P\}$ , which is located in the environmental domain  $D$ , is shown in Figure 10. In the SAH loop  $L$ , human agents  $H = \{h_1, \dots, h_M\}$  are connected to relevant features  $F$ ,  $F \subseteq D$ , via the robot swarm  $S = \{r_1, \dots, r_N\}$ . Each human agent  $h_i$  and robot agent  $r_k$  is defined by a state vector which represents, for example, cognitive states and sensor states. Each feature  $f_q$ , in turn, is defined by a state vector that describes the feature given the task objectives (i.e. the affordance, Gibson, 1979),  $f \in F$ .

Treating the swarm as a functional part of the body first means that the robot collective  $S$  implements the interface that forwards estimates of the human states  $\hat{H}$  and feature states  $\hat{F}$  via the connectome graph  $G_C = (V_C, E_C)$ , with  $V_C$  being the nodes (representing robot and human agents) and  $E_C$ ,  $E_C \subseteq V_C \times V_C$ , being the edges (chosen communication links) between a pair of agents ( $v'_C, v''_C$ ). The graph  $G_C$ , in turn, is a subset of possible connections  $G_C \subseteq G_P$ , where the agent pair ( $v'_P, v''_P$ ) is connected in  $G_P$  when they are located in their respective communication range  $R$ . Second, parts of the swarm  $P$ ,  $P \subseteq S$ , are treated as neural low-level sensory–motor units, similar to CPGs, that locally act and sense on environmental features  $F$ . Each swarm subset  $p_o \in P$  forms an observable behavioural cluster, determined by the locally similar sensed environmental feature  $\hat{f}_q$  and received human state  $\hat{H}$ , that, assuming a certain stability of  $p_o$  in time and space, can be interpreted as an artificial body part by the human observers. Both the human–feature pathway  $G_C$  and the local sensory–motor subswarm clusters  $P$  should self-organize.

Figure 11 shows the three self-organizing interaction networks of the SAH. On the individual layer, called node layer, each robot  $r_k$  has a controller that implements microlevel rules. From a neural perspective, each robot can be seen as a functionally grouped neural subpopulation (Ampatzis, Tuci, Trianni, Christensen, & Dorigo, 2009; Otte, 2018); a neural assembly (Hebb, 1949) that dictates the behaviour of the robot. The other two layers constitute the macrolevels as emerging out of the node layer's microlevel rules. The connectome layer is the ad hoc network  $G_C$ , a neural pathway that allows information exchange between human



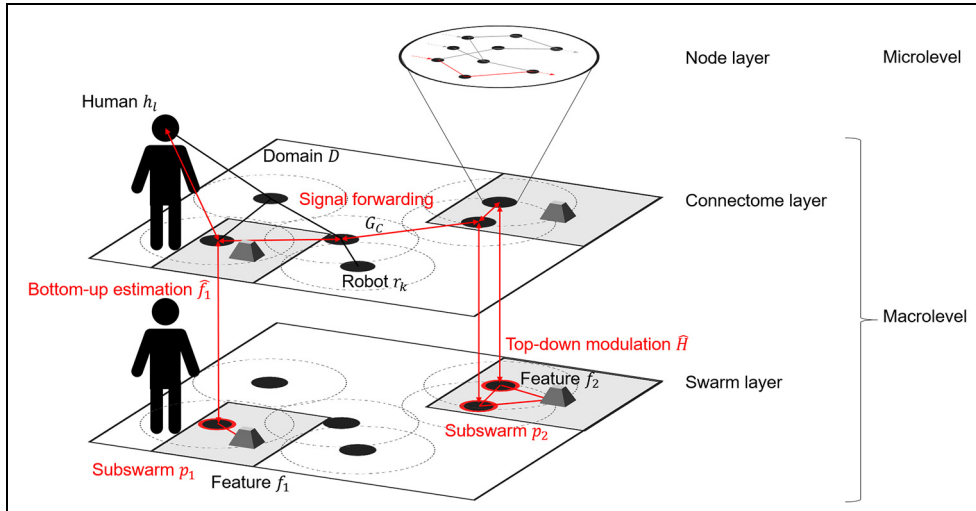
**Figure 10.** The SAH loop consists of humans  $H$  and the robot swarm  $S$ . The humans  $H$  are connected to environmental features  $F$  via the robot swarm  $S$ . The swarm  $S$  self-organizes into the connectome graph  $G_C$  and into subswarms  $P$ . The connectome graph  $G_C$ , which is a subset of the possible connections graph  $G_P$ , connects humans  $H$  with the subswarms  $P$  and forwards global human modulation  $\hat{H}$  as well as estimated feature state  $\hat{F}$ . The subswarms  $P$  emerge out of locally similar feature estimations  $\hat{f}_q$  and global human modulation  $\hat{H}$ .

operators  $H$  and environmental features  $F$ . Forming a stable  $G_C$  fundamentally relates to restricting robot movement, at least if a fast information exchange is intended, given that potential hop units  $r_k$  must be located between  $H$  and  $F$ . Local sensory–motor loops  $P$ , that also represent neural assemblies, emerge on the swarm level by an interplay of global human top–down modulation  $\hat{H}$  and respective local environmental states  $\hat{f}_q$  as perceived by individual robot–neurons  $r_k$ . The robot's controller, therefore, must implement a function  $u(\hat{H}, \hat{f}_q)$  that adjusts the behaviour of the local sensory–motor loops  $P$ .

When observed, the swarm will show diverse behaviour if the environment  $D$  is diverse, a property in line with behaviour-based robotics as the environment is thought of as being its own model (Brooks, 1991). It also makes sense from a human perspective. Imagine putting your hand over a burning candle. Surely retracting your hand (a subset of your body) will be sufficient while running away from the candle, that is, retracting your whole body, would be an unnecessary waste of energy rendering it an unintelligent decision. Taken together, the essential engineering objective is to find microlevel rules at the node layer that allow for the self-organized macrolevel formation of human–feature pathways  $G_C$  at the connectome layer and local sensory–motor loops  $P$  at the swarm layer.

Before we turn to a computational experiment that shows how to create a self-organizing graph  $G_C$  and subswarm  $P$ , we must shortly return to the human





**Figure 11.** The SAH loop consists of three self-organizing interaction networks. On the node layer, the robots' controllers specify the microlevel rules that define the macrolevel behaviours for the connectome layer and the swarm layer. The connectome layer is the robot interaction network  $G_C$  that forwards the global top-down human modulation  $\hat{H}$  and local bottom-up feature state estimations  $\hat{F}$ . Behavioural subswarms  $P$  form based on the distribution of environmental features  $F$  in the environment  $D$  and the received human modulation signals  $\hat{H}$ .

operator's perspective to illustrate how the concept of self-organization is important here as well. Assuming a successful integration of the swarm at the human operator's low-level sensory-motor control (Section 6.3), the operator will probably not be able to magically use the swarm without prior exposure. However, this is in line with the SAH, because humans also need training to learn sensory-motor patterns for their real body parts. No human baby is able to walk after birth, but with exploration, stable walking patterns form. Note how this relates to the discussion of intelligent systems that tune the robust controller by rules of adaptation (Section 3.2). Similarly, the operator should be exposed to different situations so that he is able to develop an intuition of swarm dynamics as well as how his own states influence these dynamics (Hasbach et al., 2020). This process may be described by neural plasticity (Hebb, 1949), meaning it can be seen as self-organizing on the neural level. With time, artificial body part control should feel more automatic, controllable and effortless, similar to learning new sensory-motor patterns with the natural body.

## 5. Computational experiment

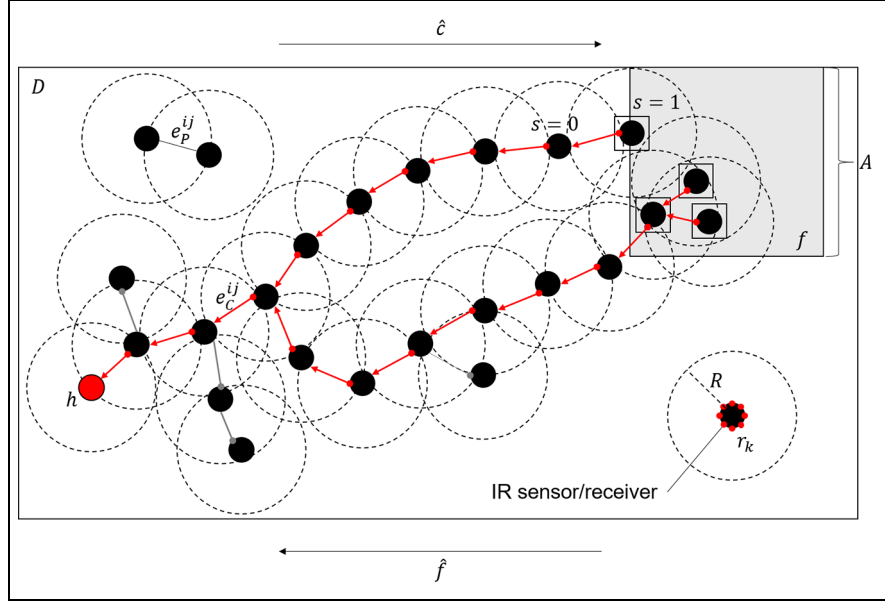
### 5.1. Problem statement

The problem is defined as finding a robot controller (node layer) that allows for the formation of the connectome graph  $G_C$  between humans  $H$  and feature set  $F$  at the connectome layer while  $F$  is coupled to the JHSL via local sensory-motor clusters  $P$  at the swarm

layer. Our objective is to demonstrate how this can be achieved by a bottom-up design, that is, combining minimal building blocks and observing what behaviours arise (Braitenberg, 1986).

Based on this objective, we were in particular interested if  $G_C$  could be found by randomly moving robots as this constitutes one of the simplest but relevant behaviours in swarm robotics. One may think of a swarm exploring an unknown environment. Extensive work in swarm robotics includes pathway formation of some sort (e.g. Campo et al., 2010; Dorigo et al., 2013; Garnier, Tache, Combe, Grimal, & Theraulaz, 2007; Hauert, Zufferey, & Floreano, 2009; Hoff, Wood, & Nagpal, 2013; Howard, Mataric, & Sukhatme, 2002; Molins, Stillman, & Hauert, 2019; Nouyan et al., 2008; Payton et al., 2001; Penders et al., 2011; Reina, Miletitch, Dorigo, & Trianni, 2015; Schmickl & Crailsheim, 2008; Sperati, Trianni, & Nolfi, 2011; Stirling, Wischmann, & Floreano, 2010; Szymanski, Breitling, Seyfried, & Wörn, 2006; Werger & Mataric, 1996). In the following, we will build on the work of Campo et al. (2010) who described a distributed algorithm for shortest path tree selection without relying on hop counts. To be free of hop counts fits to the intended property of scalability. In the work of Campo et al., robots located at a central place emitted a signal that spreads through the swarm ad hoc network like a waveform away from the central place. While they assumed that a communication link had already been established (i.e. that robots are located between a central place and a target), in the following, we show how this approach can be used to find a pathway between a





**Figure 12.** Schematic for the formation of the connectome graph  $G_C$  based on minimal principles. The human's estimated cognitive state  $\hat{c}$  is induced and forwarded in the possible connections graph  $G_P$  while each robot memorizes the incoming edge that it has first received (red and grey edges). If a robot has received a cognitive state estimate  $\hat{c}$  and has a sensor state  $s = 1$ , it induces the feature state estimate  $\hat{f}$  into the network. The feature state estimation  $\hat{f}$  travels the route back to the human (red path), based on the memorized edges of the forwarded cognitive state estimate  $\hat{c}$ . If a robot has received both  $\hat{c}$  and  $\hat{f}$ , it stops moving.

human  $H$  and a feature set  $F$  by randomly moving robots and to create subswarms  $P$ .

### 5.2. Experimental setup

The simulation was implemented in MATLAB 2019b. The robot model is assumed to be equipped with eight direction-sensitive infrared (IR) transmitter/receiver pairs (Figure 12, right bottom corner). Let the robot state be defined by

$$r_k = [x, y, \theta, d, \hat{c}, s, \hat{f}]$$

with  $x, y \in \mathbb{N}$  being the position vector in the two-dimensional domain  $D$ ,  $\theta$  being the orientation of the robot ( $\theta \in \{1, \dots, 8\}$ ),  $d$  being the direction of a received forwarded signal based on the IR receiver locations that is simplified as the corresponding edge of the graph  $G_P$  ( $d \in E_P$ ),  $\hat{c}$  being the received estimate of the human's cognitive state,  $s$  being the sensor state measuring the environmental feature  $F$ , and  $\hat{f}$  being the received sensor signal of  $f_q$  that was forwarded by another robot ( $\hat{c}, s, \hat{f} \in \{0, 1\}$ ).

The state of the human is defined by

$$h = [x, y, c]$$

with  $c \in \{0, 1\}$  being a particular cognitive state (e.g. arousal) of the human that is either active or not.

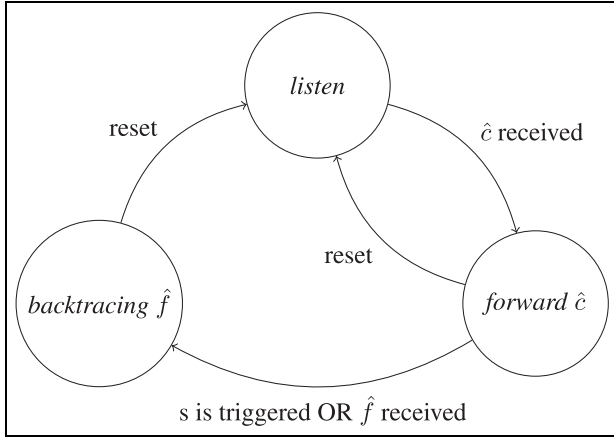
Each simulation run was repeated over 1000 samples over a range of different swarm sizes  $N = 100 \dots 200$ .

During initiation of each run, one human agent  $h$  and a swarm  $S$  of  $N$  robots were placed as point masses in a  $100 \times 100$  grid domain,  $D \subseteq \mathbb{N}^2$ . The human cognitive state  $c$  was set to 1 (active) while the received states of the robots  $\hat{c}$  were set to 0 (inactive). The locations  $[x, y]$  as well as the robot orientations  $\theta$  were chosen randomly.  $D$  included a randomly placed feature area that is defined by a square with  $A = 20$  being the side length.

While the position of the human agent  $h$  was fixed during a simulation run, the robots wandered randomly. At each iteration of the simulation loop, a robot would choose a new orientation  $\theta$  with equal probability when facing another agent or the grid boundary in the cell in front of it while moving straight forward if no collision was imminent. The robot's feature sensor state was set to  $s = 0$  if a robot was outside the feature area  $f$  or set to  $s = 1$  if it was inside. Based on the robots' position at each iteration, the possible connections were constructed by comparing the current euclidean distance with the robots' communication range  $R = 10$ ,  $\{e_{ij}^{ij} \in E_P : \sqrt{(x^i - x^j)^2 + (y^i - y^j)^2} \leq R\}$  (Figure 12). The  $G_P$  was then used to establish the connectome graph  $G_C$  via the microlevel rules described in the following.

### 5.3. Connectome layer

The human agent blindly transmits his cognitive state  $c$  to all robots in his communication range (Figure 12). Figure 13 shows the robot controller as a finite-



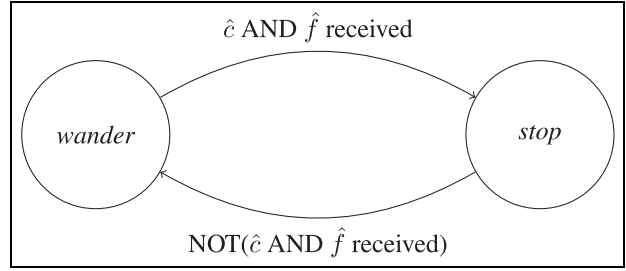
**Figure 13.** FSA for local signal forwarding.

state-automaton (FSA) describing the signal forwarding based on  $G_P$ . The controller has three behavioural states. At each iteration of the simulation loop, a reset signal first renders all robots into the state *listen*. This may be implemented by pre-programmed time intervals that reset robots. During *listen*, the robot waits for a reception of an estimated cognitive state  $\hat{c}$  from an arbitrary source that can either be the human agent or another robot. After reception of  $\hat{c}$ , the direction  $d$  of the incoming signal is saved and the state *forward  $\hat{c}$*  is activated. In *forward  $\hat{c}$* , the robot forwards the received  $\hat{c}$  based on  $G_P$ . Similar to the work by Campo et al. (2010), robots that are not in the *listen* state do not accept any newly incoming  $\hat{c}$  transmissions. This ensures that the human state  $\hat{c}$  spreads radially away from the human agent with minimal node jumps (Figure 12), given the assumption that more node jumps take more time to arrive at a target node. If a robot receives a  $\hat{c}$  and has sensed the environmental feature via sensor  $s$  or has received the estimated feature  $\hat{f}$  via another robot, it means that a path between human  $h$  and  $f$  has been found. During *backtracing  $\hat{f}$* , the robot emits the estimated  $\hat{f}$  only in the direction of  $d$ . The signal  $\hat{f}$ , therefore, travels back to the  $h$  via the route of the  $\hat{c}$  spreading (red paths in Figure 12).

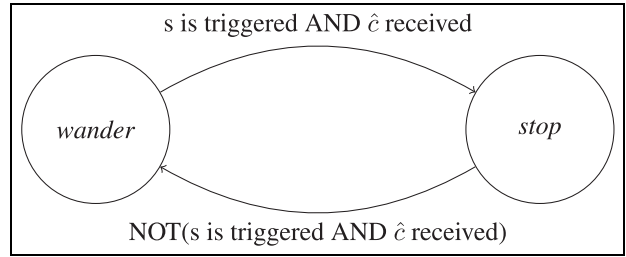
The emergence of a stable connectome graph  $G_C$  over time now occurs via a simple rule; if a robot has received both  $\hat{c}$  and  $\hat{f}$ , it must be a node of the  $G_C$  and, therefore, stops moving (Figure 14). The restriction of robot movement to form pathways via a swarm was also applied by Nouyan et al. (2008) and Molins et al. (2019).

#### 5.4. Swarm layer

Based on the forwarded signals, the robot controller for the emergence of the local subswarms  $P$  is straight forward (Figure 15). If a robot has received the  $\hat{c}$  and is close to the feature  $f$  as measured by  $s$ , a change in



**Figure 14.** FSA for pathway formation.

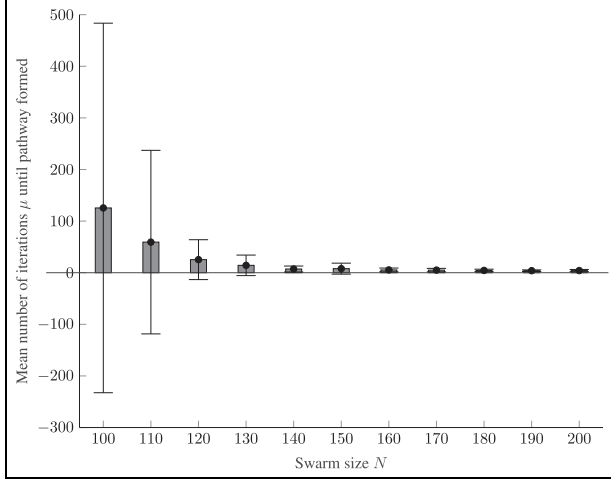


**Figure 15.** FSA for local sensory-motor loop formation.

robot behaviour is triggered. Here, we simply stopped the robot to show the formation of  $P$  as clusters, but the change in state could relate to a parameter change of a swarm algorithm as well as to a switch between swarm algorithms (Kolling et al., 2015). One may think of shifting from walking to running locomotion based on the adjustment of CPG inputs. Note however that some robots at the feature  $f$  must be gateway nodes and should, therefore, be restricted in their movement so that  $G_C$  is not disrupted.

#### 5.5. Results

Figure 16 shows the performance of the algorithm in terms of required iterations until a pathway between  $h$  and  $f$  has formed over different swarm sizes  $N$ . With a swarm size of  $N = 100$ , it took considerable time until a pathway would emerge with a mean step time  $\mu_{N=100} = 125.5$ . A standard deviation of  $\sigma_{N=100} = 358$  showed the sensitivity of the algorithm to different spatial arrangements between human and feature. With swarm sizes lower than around  $N = 100$ , we found that pathways did not always emerge in acceptable time. Figure 16 shows that the formation of the pathway  $G_C$  depends on the size of the swarm  $N$  with  $\mu_{N=200} = 4.18$  and  $\sigma_{N=200} = 2.14$ . This result is consistent with the work of Nouyan et al. (2008) and is not surprising as the robots wander randomly. Larger swarm sizes increase the number of possible connections, leading to a higher chance to find a  $G_C$ . Importantly, provided there is a high swarm density in a restricted area, it can be seen how even simple micro-level rules combined with random robot movement



**Figure 16.** Mean number of simulation loop iterations  $\mu$  until pathway has formed over different swarm sizes  $N$ . With an increasing number of swarm robots, a pathway between human  $h$  and feature  $f$  is found faster.

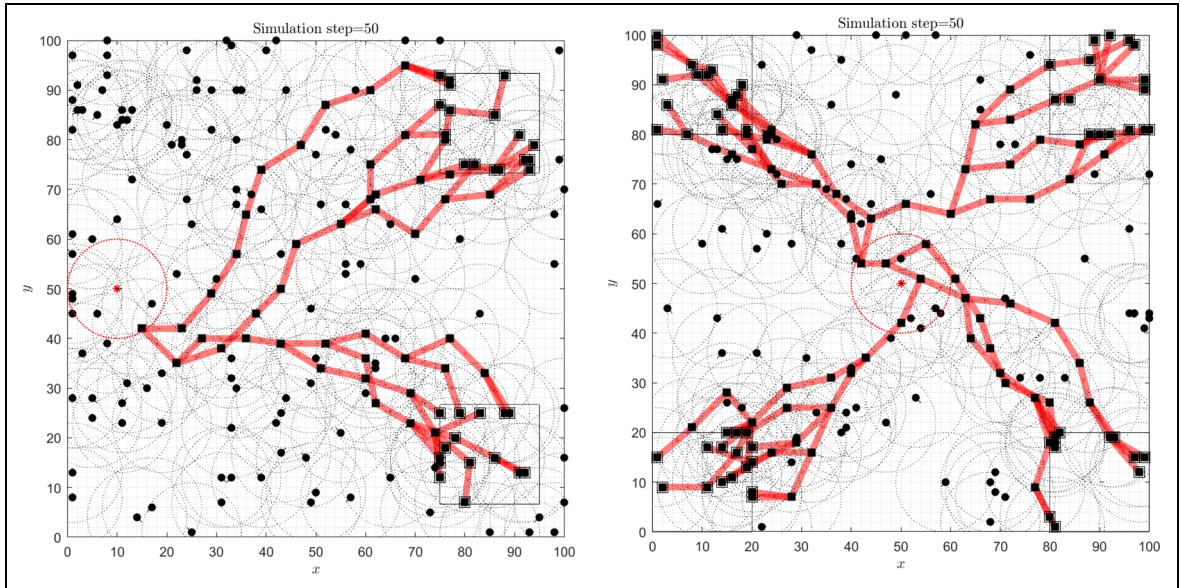
result in the self-organization of a  $G_C$ . No positive feedback mechanism is required.

An additional property of the algorithm is the shortening of the path with simulation loop iterations. As the robots regularly reset to form a new  $G_C$ , the random wandering of robots will occasionally find a solution with fewer node jumps. In addition, the design is not restricted to a single feature  $f_q$ . Figure 17 shows two examples for path formation with multiple features. By the same microlevel rules, multiple pathways with redundant parallelized connections have formed between the features and the human. The figure also

shows the self-organization of multiple subswarm clusters  $p_o$  based on the spatial distribution of the environmental features.

The algorithm can be evaluated in terms of human-swarm intelligence (Section 3.2). The performance is robust in that it will find a pathway (a subgoal of the SAH) independently of the spatial arrangement of  $h$  and  $f$  (the situation set), even though it may take longer if  $h$  and  $f$  are further apart from each other. Given a certain swarm density, where the exact minimum number of required robots depends on multiple parameters such as the nature of  $D$  and the robots' communication range  $R$ , it is scalable in the sense that an increased swarm size will provide a higher chance for a  $G_C$  to be identified. Scalability in terms of multiple humans has not been evaluated. The evaluation of adaptation depends on the definition of the internal organization. The macrolevel, where the connectome graph  $G_C$  and the swarm clusters  $P$  represent the internal organization, shows adaptivity since the network structure of the swarm changes over time to fit the current situation. However, the algorithm is not adaptive on the microlevel (node layer) as the robots' controllers are not changed. This is in accordance with the SAH, which allocates the responsibility for adaptation of the control model to the human operator (Section 4.3).

Based on the formation of  $G_C$  and  $P$ , now imagine an aversive stimuli being present at  $p_1$  (cf. Mathews et al., 2017). By simple behaviour-based control, we can define that the robots forming the artificial body part  $p_1$  retract from the stimuli. Now say that the human state estimate  $\hat{c}$  signals high arousal. We define that the threshold to react to stimuli is some positive linear



**Figure 17.** Simulation run examples with  $N = 200$  robots, showing the formation of swarm clusters  $P$  and redundant pathways in  $G_C$ . Left: two feature areas. Right: four feature areas.

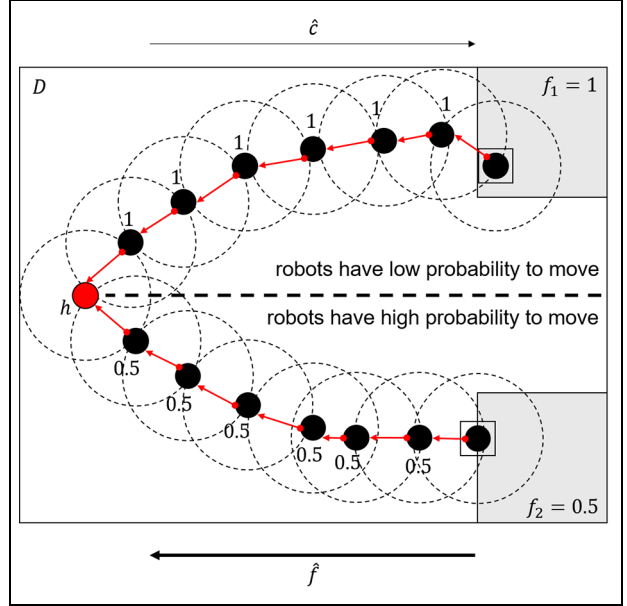
function of the estimated arousal  $u(\hat{c})$ . This will result in a decreased retraction from the stimuli in  $p_1$ . It can be said that top-down influence modulated low-order stereotypical behaviour, similar to the increase in swarm alertness based on Tom's distress. Note again how the SAH is in line with behaviour-based approaches such as the subsumption architecture (Arkin, 1998; Brooks, 1986).

## 6. Advanced implications

### 6.1. On neural computation in SAH

The robot controllers described above resulted in the generation of a pathway between human  $h$  and feature set  $F$ . This pathway formation relates to the genetically guided establishment of neural connections (Braitenberg & Schüz, 1991; Sanes & Jessell, 2013b). While the above computational experiment focused on simplicity, a more biologically plausible and more efficient algorithm may be a searching robot chain that originates from the human (Nouyan et al., 2008), similar to growing biological axons that are guided by molecular cues (Sanes & Jessell, 2013b).

The established connections are then tuned to fit the dynamic environment (Sanes & Jessell, 2013a). From the outside, this tuning is observed as learning. Donald Hebb (1949) has described a process for the tuning of connections on the neural level. Hebbian learning is often summarized as 'neurons that fire together, wire together', although it should be remarked that this statement alone is too simplified. We may implement Hebbian learning in the  $G_C$  by adjusting the *stop* behaviour of the FSA in Figure 14. Instead of stopping if a robot receives both  $\hat{c}$  and  $\hat{f}_q$ , a function that relates the value of  $\hat{f}_q$  to a stopping probability is implemented. Figure 18 shows an example containing two features  $f_1$  and  $f_2$  placed at an equal distance from the human  $h$  while  $f_1$  has a value of 1 and  $f_2$  has a value of 0.5. The activity related to  $\hat{f}_q$  is spread through the robot-neuron connectome  $G_C$ . If the robot-neurons fire a lot together, the pathway is rendered importantly by Hebbian learning leading to a strengthening of the connections by constraining robot movement to stabilize the pathway. However, with less forwarded activity, the pathway may not be as important; robots start to move again, which potentially leads to a disruption of the specific pathway. We may say that the connection is forgotten to free robot-neuron resources so that new important pathways may be learned. Probability triggers for movement control to establish stable and unstable pathways have been used in swarm robotics (Nouyan et al., 2008). Alternatively, Hebbian learning may be simulated by adjusting the distance between the robots (Monaco et al., 2020) or by attaching virtual synaptic weights to the edges in  $G_C$  (Campo et al., 2010; Otte, 2018).



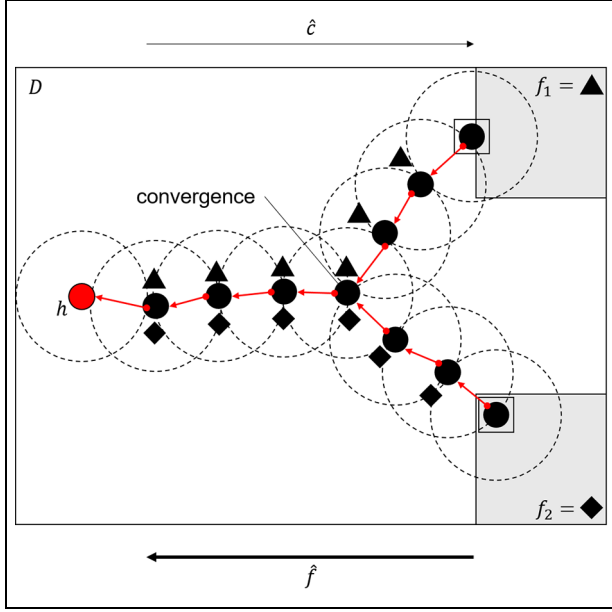
**Figure 18.** In the SAH, synaptic tuning may be implemented by restricting robot movement based on the importance of the pathway.

We now turn to the topic of providing feedback via the swarm. In Section 4.4, it was discussed that feedback should be provided at the low-level sensory system. The localization of a feature source, a global information, may then be achieved by an actively exploring human agent moving through a swarm gradient. This swarm gradient may be seen as the functional equivalent of neural population coding (Rolls, 2016). It may also be asked how neural networks process local sensory information to produce global knowledge. Known to practitioners familiar with neural networks is the notion of convergence, where the receptive field of neurons are broadened over neural layers (Hubel & Wiesel, 1962; Rolls, 2016). Figure 19 shows convergence in the SAH. Information from two different features  $f_1$  and  $f_2$  are bound together by robot-neurons receiving inputs from both features. These nodes are candidates for implementing receptive fields that are tuned to both features. By convergence, global information about the environment can be estimated by the swarm connectome and provided to the human, similar to sensory upstreams in the biological nervous system. For example, Tom may be signalled by his wearable interface that both fire source and a victim are located together, compared to a victim at another location without a fire source. This helps Tom decide who to rescue first. The concept of convergence was already applied in swarm robotics (Mathews et al., 2017).

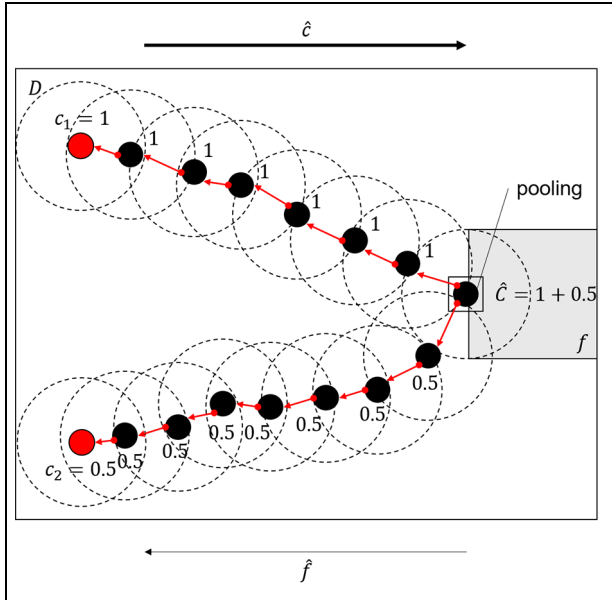
### 6.2. On human scalability in SAH

In Section 3.2, scalability was defined as a desired capability for HSI. In addition to swarm scalability as





**Figure 19.** In the SAH, neural convergence may be implemented by robots that are connected to multiple features.



**Figure 20.** In the SAH, updating the behaviour of local sensory-motor loops can be based on the cognitive states of multiple humans.

described, one may adjust the controllers above to include multiple human agents. Figure 20 shows an example of two pathways between one feature  $f$  and two humans  $h_1$  and  $h_2$ . Robots at the local sensory-motor loop implement the FSA of Figure 15, with the difference that in order to *stop*,  $\hat{c}_1$  and  $\hat{c}_2$  are pooled before given to the function  $u$  that adjusts robot behaviour. Pooling of neighbouring opinions is a common approach in swarm robotics (Hamann, 2018).

The cognitive states of humans may also influence each other via the swarm pathway, resulting in a ‘socio-technical group mind’ established by the self-organizing swarm. Note how the ‘socio-technical group mind’ relates to human swarm intelligence (Tavakoli et al., 2016) as humans displaying swarm intelligence. This ‘group mind’ may be used to combine proximal and remote interaction where the states of humans with local situational awareness (narrow, detailed information) are pooled with the states of humans that have global situational awareness (broad, abstract information). The ‘group mind’ highlights the view of human scalability from the neural perspective; human operators can be seen as different cognitive modules that collectively vote for JHSL output (Minsky, 1988; Richards, 2015).

### 6.3. On artificial body parts in SAH

In the SAH, the swarm is integrated into the human’s nervous system function. We have said that this fundamentally relates to passive interaction and to providing bottom-up feedback (Section 4.4). In the following paragraphs, additional considerations are discussed.

Joining nervous system and swarm into one system means that the state of one depends, at least in part, on the other. There are two routes to establish a communication between human and nervous system with the swarm: neural and behavioural. A method is categorized as neural if it is directly targeted at the biological nervous system to implement a translation between neural state and swarm state. Neural methods, therefore, include both invasive and non-invasive brain-computer interfaces (Karavas et al., 2017; Lebedev & Nicolelis, 2017; Wolpaw, Birbaumer, McFarland, Pfurtscheller, & Vaughan, 2002). On the contrary, behavioural methods refer to interfacing with an additional transformation between nervous system and swarm. Thus, one infers and manipulates the neural states based on psychological and psychophysiological states (Alboul et al., 2008; Schwarz & Fuchs, 2017) that are, at least partly, a result of neural computation. For example, transcranial magnetic stimulation (TMS) application at the visual cortex would directly target neural states, while providing simple visual cues ‘hitch-hike’ a biological sensor that is connected to the nervous system and by this route indirectly influence the neural state.

It will, of course, not be sufficient to establish a mere state dependence of nervous system and swarm to produce a real SAH. After all, traditional human-machine interfaces establish connections via the behavioural routes (e.g. clicking a mouse button is a behavioural state that is an output of nervous system control). To include the swarm in the human’s nervous system, the brain must be tricked into integrating the swarm into its body representation. This is achieved if the human operator has the illusion that the swarm feels like part

of himself, if it feels like an artificial body part. This requires first, and most fundamentally, that the subswarms form operational unities in the sense that they create, recreate and hold into an observable form separated from the environment (Varela, Maturana, & Uribe, 1974). The experiment above has shown how this can be in principle done with minimal effort. Second, additional requirements must be formulated that satisfy how the brain integrates and controls the body. This renders contemporary work on the brain's body representation, tool use and neural engineering relevant for the SAH (Abbott, 2006; Botvinick & Cohen, 1998; Braun et al., 2018; Cardinali et al., 2009; Cheng, Ehrlich, Lebedev, & Nicoletis, 2020; Kieliba et al., 2021; Kwakkel, Kollen, & Krebs, 2008; Lebedev & Nicoletis, 2017; Maravita & Iriki, 2004; Miller et al., 2018; Schettler, Raja, & Anderson, 2019; Seth, 2013). Work on body ownership and agency suggest that bottom-up multi-modal feedback should be spatio-temporally correlating and that the subswarms probably need to look like part of the body, for example, a hand (Braun et al., 2018). Robot swarms, especially shape-changing swarms (Mathews et al., 2017; Oh et al., 2017; Rubenstein et al., 2014; Slavkov et al., 2018) that are observed through augmented reality displays (Payton et al., 2004; Reina et al., 2017), can therefore be advanced engineering test-beds for theories on how the brain controls the body.

In sum, the SAH requires the translation of neural signals to control inputs for self-organized body parts in accordance with body control. The indirect route may already be a feasible solution for real-world applications, for example, by utilizing the human's position (Alboul et al., 2008) and physiological measurements (Schwarz & Fuchs, 2017) as control inputs. For example, fire-fighter Tom (Section 1) rapidly changing into a defensive body posture can be used as a control signal for the swarm body parts engaging in a reflex-like protection of Tom. However, this behavioural classification seems generally limited to low-resolution and context-dependent interpretations of the neural state. Brain-computer interface technology has the potential of high-resolution classification, although the technology is not yet mature. In the future, Tom may be equipped with an advanced brain-computer interface technology that provides high-resolution classifications of cognitive states and adapted sensory-motor signals. This could allow him, for example, to feel the location (i.e. the direction and the distance) of a victim, alike to having a sixth sense, while allocating more robots to the estimated location of the victim, similar to closing your hands around an object.

## 7. Conclusion

This work started with a description of JHSL components, that is, human, swarm and interface. The design

of a good JHSL must inevitably take into account the properties of the components in relation to each other, in particular to consider the centralized-decentralized control and communication trade-off. It was stated that a systems theory in terms of a design theory, that is, a theory that predicts what would be a good combination of human, swarm and interface, can be beneficial in this endeavour. Based on the discussion of desired JHSL properties, it was concluded that a design theory must formulate solutions for how the centralized control demands of human operators and the decentralized control demands of swarms can be joined into one goal-seeking system while promoting sustained adaptive, robust and scalable behaviour. This was referred to as the challenge of human-swarm intelligence.

The design theory that we called the SAH, the swarm-amplified human, predicts that integrating semi-autonomous robot swarm clusters into the human's low-level nervous system is a good design solution. It was argued that doing so establishes a capable joint-systems controller that can react fast to local changes but can also deal with more complex scenarios while balancing between central and distributed control. JHSL adaptivity is mainly supported by amplifying the human operator's natural cognitive capabilities by extending his sensory-motor possibilities while robustness and scalability are preserved by treating the swarm as a neural machinery.

Subsequent discussion focused on relating work in swarm robotics and theoretical neuroscience when both systems are modelled as self-organizing systems. The SAH defines three levels of self-organization: the node layer, the connectome layer and the swarm layer. The node layer defines the microlevel interactions between robot-neurons that result in the formation of the swarm as interface between human operators and environmental features as well as swarm clusters that locally act on the environmental features and constitute the artificial body parts.

In sum, the SAH defines the following set of four central interdependent requirements to join the perspectives of human, swarm and interface. These requirements should be seen as hypotheses that must be subjected to empirical probing through testing prototypes that preferably satisfy all four requirements.

1. Rather than controlling the whole swarm, the swarm should self-organize into semi-autonomous subswarms that locally interact with relevant features of the environment while being modulated by the human's cognitive state. This mimics the abstract neural control architecture.
2. Rather than being the target of the interaction, the swarm should self-organize as the interface between the human operator and relevant features of the environment. This mimics the formation of neural



pathways in the nervous system that are tuned to relevant environmental stimuli.

3. Rather than consciously supervising a robot swarm, the human state should be translated into swarm commands in a passive yet controllable manner with the option to switch to manual control of subswarms. This mimics the daily life phenomenology of body control.
4. Rather than providing artificially constructed feedback, environmental feedback should be provided in a bottom-up manner featuring raw stimulation of biological sensory systems. This mimics the dependence of cognition on lower-order sensory-motor capabilities.

The implemented computational experiment, which focused on the first two of the requirements above, showed how the human-feature pathways and local subswarms can self-organize with minimal logic. It was further discussed how the SAH theory provides a range of additional solutions for HSI design that blur the distinction between swarm intelligence and neural computation. This is to be desired because a good theory should encompass, or relate to, a high number of (possible) observations.

We may say that the described robot controllers in the computational experiment in itself are rather simple and, therefore, perhaps of little value. Such a proposition, however, would miss the points we want to make tremendously. First, self-organization is formalized as equipping a system with complex behaviours even if the mechanism is simple. To quote Valentino Braitenberg discussing his law of uphill analysis and downhill invention: *'A psychological consequence of this is the following: when we analyse a mechanism, we tend to overestimate its complexity'* (Braitenberg, 1986, p. 20). Part of this work aims at showing how even simple logic can be utilized for the design of HSI, similar to the approach adapted in swarm engineering and complexity science. Second, this work is on how to best design HSI from a global point of view. It may, therefore, be categorized as a work of system-theoretical cybernetics, defined as *'the framework in which all individual [systems] may be ordered, related and understood'* (Ashby, 1961a, p. 2). One may add to this *'...and joined'*, an implicit assumption applied in human-machine systems engineering (Hollnagel & Woods, 2005; Licklider, 1960; Woods & Hollnagel, 2006). In essence, a given design theory is a cybernetic model describing how different systems may be joined into one new system with superior capabilities.

Starting at the top of our argument, the SAH as design theory allows predictions that we deduced as a testable prototype. We have tested the algorithm for robustness and scalability in the context of the SAH and concluded that it fits demands. One may, therefore,

proceed with caution for we have only tested an isolated part of the SAH theory, in accepting the SAH as a suitable candidate to inform design.

The application of the empirical cycle that includes global theories for the big picture is essential in the scientific design of artificial systems such as HSI. Unfortunately, the benefit of understanding the design space is often overlooked in the pursuit of producing artificial systems fast and cheap. Based on a good design theory, design patterns can be constructed that fit to particular environmental demands and mission objectives (Alexander, 1977; Woods & Hollnagel, 2006) such as Tom searching for survivors. In the long run, formulating design theories is, therefore, cheaper because the impact of design decisions on the overall system can be predicted more reliable.

In general, the scientifically oriented JHSL-engineer takes the role of an external adaptation controller who updates the internal organization of the JHSL to fit expected situations (Woods & Hollnagel, 2006). Design is a process of intelligence after all, thus the selection of appropriate actions (Simon, 1982). We hope that our work shows that treating robot swarms as neural machines can inform the selection of design actions in HSI. From the view of natural science, the construction of artificial distributed intelligence that merges neural and swarm computation provides a method for comparing and integrating theories about nervous systems and swarms including whether evolutionary pressure could indeed have led to the invention of similar computational underpinnings across distributed systems observed in nature.

## Acknowledgments

The authors thank the department of Human-Machine-Systems for providing resources and support.


## Declaration of conflicting interests

The author(s) declared no potential conflicts of interest with respect to the research, authorship and/or publication of this article.

## Funding

The author(s) received no financial support for the research, authorship and/or publication of this article.

## ORCID iD

Jonas D Hasbach  <https://orcid.org/0000-0002-8104-3453>

## Note

1. The neuroscientist may forgive us for referring also to subcortical functions as cognitive for the sake of convenience.

## References

- Abbott, A. (2006). In search of the sixth sense. *Nature*, 442(7099), 125–127.
- Alboul, L., Saez-Pons, J., & Penders, J. (2008). Mixed human-robot team navigation in the GUARDIANS project. In *2008 IEEE International Workshop on Safety, Security and Rescue Robotics* (pp. 95–101). IEEE.
- Alexander, C. (1977). *A pattern language: Towns, buildings, construction*. Oxford University Press.
- Alonso-Mora, J., Lohaus, S. H., Leemann, P., Siegwart, R., & Beardsley, P. (2015). Gesture based human multi-robot swarm interaction and its application to an interactive display. In *2015 IEEE International Conference on Robotics and Automation (ICRA)* (pp. 5948–5953). IEEE.
- Ampatzis, C., Tuci, E., Trianni, V., Christensen, A. L., & Dorigo, M. (2009). Evolving self-assembly in autonomous homogeneous robots: Experiments with two physical robots. *Artificial Life*, 15(4), 465–484.
- Arbib, M. A. (1989). *The metaphorical brain 2: Neural networks and beyond*. John Wiley & Sons.
- Arkin, R. C. (1998). *Behavior-based robotics* (1st ed.). The MIT Press.
- Ashby, R. W. (1962). Principles of the self-organizing system. In H. von Foerster & G. Zopf (Eds.), *Principles of self-organization: Transactions of the University of Illinois Symposium* (pp. 225–278). Pergamon Press.
- Ashby, W. R. (1947). Principles of the self-organizing dynamic system. *Journal of General Psychology*, 37, 125–128.
- Ashby, W. R. (1960). *Design for a brain: The origin of adaptive behaviour* (2nd ed.). Chapman & Hall.
- Ashby, W. R. (1961a). *An introduction to cybernetics*. Chapman & Hall.
- Ashby, W. R. (1961b, May 9–11). *What is an intelligent machine?* [Conference session]. *Western Joint IRE-AIEE-ACM Computer Conference*, Los Angeles, California, USA.
- Ashcraft, C. C., Goodrich, M. A., & Crandall, J. W. (2019). Moderating operator influence in human-swarm systems. In *2019 IEEE International Conference on Systems, Man and Cybernetics (SMC)* (pp. 4275–4282). IEEE.
- Bainbridge, L. (1983). Ironies of automation. *Automatica*, 19, 775–779.
- Balch, T., & Parker, L. E. (2002). *Robot teams: From diversity to polymorphism*. CRC Press.
- Bales, G., & Kong, Z. (2017). Neurophysiological and behavioral studies of human-swarm interaction tasks. In *2017 IEEE International Conference on Systems, Man, and Cybernetics (SMC)* (pp. 671–676). IEEE.
- Barca, J. C., & Sekercioglu, Y. A. (2013). Swarm robotics reviewed. *Robotica*, 31(3), 345–359.
- Bashyal, S., & Venayagamoorthy, G. K. (2008). Human swarm interaction for radiation source search and localization. In *2008 IEEE Swarm Intelligence Symposium* (pp. 1–8). IEEE.
- Bayindir, L. (2016). A review of swarm robotics tasks. *Neurocomputing*, 172, 292–321.
- Beer, S. (1985). *Diagnosing the system for organizations*. John Wiley & Sons.
- Birattari, M., Ligt, A., Bozhinoski, D., Brambilla, M., Francesca, G., Garattoni, L., Garzón Ramos, D., Hasselmann, K., Kegeleirs, M., & Kuckling, J. (2019). Automatic off-line design of robot swarms: A manifesto. *Frontiers in Robotics and AI*, 6, Article 59.
- Bonabeau, E., Dorigo, M., Marco, D. d. R. D. F., Theraulaz, G., & Theraulaz, G. (1999). *Swarm intelligence: From natural to artificial systems* (No. 1). Oxford University Press.
- Botvinick, M., & Cohen, J. (1998). Rubber hands ‘feel’ touch that eyes see. *Nature*, 391(6669), 756.
- Braitenberg, V. (1986). *Vehicles: Experiments in synthetic psychology*. MIT Press.
- Braitenberg, V., & Schüz, A. (1991). *Anatomy of the cortex: Statistics and geometry*. Springer.
- Brambilla, M., Ferrante, E., Birattari, M., & Dorigo, M. (2013). Swarm robotics: A review from the swarm engineering perspective. *Swarm Intelligence*, 7(1), 1–41.
- Braun, N., Debener, S., Sychala, N., Bongartz, E., Sörös, P., Müller, H. H. O., & Philipsen, A. (2018). The senses of agency and ownership: A review. *Frontiers in Psychology*, 9, Article 535.
- Brooks, R. A. (1986). A robust layered control system for a mobile robot. *IEEE Journal on Robotics and Automation*, 2(1), 14–23.
- Brooks, R. A. (1990). Elephants don’t play chess. *Robotics and Autonomous Systems*, 6(1–2), 3–15.
- Brooks, R. A. (1991). *Intelligence without reason*. Massachusetts Institute of Technology.
- Brown, D. S., Jung, S.-Y., & Goodrich, M. A. (2014). Balancing human and inter-agent influences for shared control of bio-inspired collectives. In *2014 IEEE International Conference on Systems, Man, and Cybernetics (SMC)* (pp. 4123–4128). IEEE.
- Camazine, S., Deneubourg, J.-L., Franks, N. R., Sneyd, J., Theraulaz, G., & Bonabeau, E. (2003). *Self-organisation in biological systems*. Princeton University Press.
- Campo, A., Gutiérrez, Á., Nouyan, S., Pinciroli, C., Longchamp, V., Garnier, S., & Dorigo, M. (2010). Artificial pheromone for path selection by a foraging swarm of robots. *Biological Cybernetics*, 103(5), 339–352.
- Cardinali, L., Frassinetti, F., Brozzoli, C., Urquizar, C., Roy, A. C., & Farnè, A. (2009). Tool-use induces morphological updating of the body schema. *Current Biology*, 19(12), R478–R479.
- Carrillo-Zapata, D., Milner, E., Hird, J., Tzoumas, G., Varadanega, P. J., Sooriyabandara, M., Giuliani, M., Winfield, A. F. T., & Hauert, S. (2020). Mutual shaping in swarm robotics: User studies in fire and rescue, storage organization, and bridge inspection. *Frontiers in Robotics and AI*, 7, Article 53.
- Chen, J. Y. C., & Barnes, M. J. (2014). Human-agent teaming for multirobot control: A review of human factors issues. *IEEE Transactions on Human-Machine Systems*, 44(1), 13–29.
- Cheng, G., Ehrlich, S. K., Lebedev, M., & Nicoletis, M. A. L. (2020). Neuroengineering challenges of fusing robotics and neuroscience. *Science Robotics*, 5(49), Article eabd1911.
- Clark, A., & Chalmers, D. (1998). The extended mind. *Analysis*, 58(1), 7–19.
- Clynes, M. E., & Kline, N. S. (1960). Cyborgs and space. *Astronautics*, 5, 26–27.
- Conant, R. C. (1976). Laws of information which govern systems. *IEEE Transactions on Systems, Man, and Cybernetics*, 4, 240–255.

- Conant, R. C. (1981). *Mechanisms of intelligence: Ashby's writings on cybernetics*. Intersystems Publications.
- Conant, R. C., & Ashby, W. R. (1970). Every good regulator of a system must be a model of that system. *International Journal of Systems Science*, 1(2), 89–97.
- Cooke, N. J., Gorman, J. C., Myers, C. W., & Duran, J. L. (2013). Interactive team cognition. *Cognitive Science*, 37(2), 255–285.
- Coppin, G., & Legras, F. (2011). Autonomy spectrum and performance perception issues in swarm supervisory control. *Proceedings of the IEEE*, 100(3), 590–603.
- Coucke, N., Heinrich, M. K., Cleeremans, A., & Dorigo, M. (2020). Hugos: A multi-user virtual environment for studying human–human swarm intelligence. In M. Dorigo, T. Stützle, M. J. Blesa, C. Blum, H. Hamann, M. K. Heinrich, & V. Strobel (Eds.), *International Conference on Swarm Intelligence* (pp. 161–175). Springer.
- Couzin, I. D. (2009). Collective cognition in animal groups. *Trends in Cognitive Sciences*, 13(1), 36–43.
- Crandall, J. W., Anderson, N., Ashcraft, C., Grosh, J., Henderson, J., McClellan, J., Neupane, A., & Goodrich, M. A. (2017). Human-swarm interaction as shared control: Achieving flexible fault-tolerant systems. In D. Harris (Ed.), *International Conference on Engineering Psychology and Cognitive Ergonomics* (pp. 266–284). Springer.
- Dawkins, R. (1976). *The selfish gene*. Oxford University Press.
- Dorigo, M., Floreano, D., Gambardella, L. M., Mondada, F., Nolfi, S., Baaboura, T., Birattari, M., Bonani, M., Brambilla, M., & Brutschy, A. (2013). Swarmanoid: A novel concept for the study of heterogeneous robotic swarms. *IEEE Robotics & Automation Magazine*, 20(4), 60–71.
- Dorigo, M., Theraulaz, G., & Trianni, V. (2020). Reflections on the future of swarm robotics. *Science Robotics*, 5(49), Article eabe4385.
- Doyle, J. C., & Csete, M. (2011). Architecture, constraints, and behavior. *Proceedings of the National Academy of Sciences*, 108(Suppl. 3), 15624–15630.
- Dudek, G., Jenkin, M. R. M., Milios, E., & Wilkes, D. (1996). A taxonomy for multi-agent robotics. *Autonomous Robots*, 3(4), 375–397.
- Ellwart, T. (2011). Assessing coordination in human groups: Concepts and methods. In M. Kolbe, M. Boos, T. Ellwart, & P. M. Kappeler (Eds.), *Coordination in human and primate groups* (pp. 119–135). Springer.
- Ellwart, T., Happ, C., Gurtner, A., & Rack, O. (2015). Managing information overload in virtual teams: Effects of a structured online team adaptation on cognition and performance. *European Journal of Work and Organizational Psychology*, 24(5), 812–826.
- Friston, K. (2010). The free-energy principle: A unified brain theory? *Nature Reviews Neuroscience*, 11(2), 127–138.
- Garnier, S., Tache, F., Combe, M., Grimal, A., & Theraulaz, G. (2007). Alice in pheromone land: An experimental setup for the study of ant-like robots. In *2007 IEEE swarm intelligence symposium* (pp. 37–44). IEEE.
- Gibson, J. J. (1979). *The ecological approach to visual perception*. Houghton-Mifflin.
- Gildert, N., Millard, A. G., Pomfret, A., & Timmis, J. (2018). The need for combining implicit and explicit communication in cooperative robotic systems. *Frontiers in Robotics and AI*, 5, Article 65.
- Goodrich, M. A., & Schultz, A. C. (2008). *Human-robot interaction: A survey*. Now Publishers.
- Gromov, B., Gambardella, L. M., & Di Caro, G. A. (2016). Wearable multi-modal interface for human multi-robot interaction. In K. Melo (Ed.), *2016 IEEE International Symposium on Safety, Security, and Rescue Robotics (SSRR)* (pp. 240–245). IEEE.
- Grosh, J. R., & Goodrich, M. A. (2020). Multi-human management of robotic swarms. In M. Kurosu (Ed.), *International Conference on Human-Computer Interaction* (pp. 603–619). Springer.
- Gross, R., & Dorigo, M. (2009). Towards group transport by swarms of robots. *International Journal of Bio-Inspired Computation*, 1(1–2), 1–13.
- Hamann, H. (2018). *Swarm robotics: A formal approach*. Springer.
- Hasbach, J. D., Witte, T. E. F., & Bennewitz, M. (2020). On the importance of adaptive operator training in human-swarm interaction. In R. A. Sottilare, & J. Schwar (Eds.), *International Conference on Human-Computer Interaction* (pp. 311–329). Springer.
- Hauert, S., Zufferey, J.-C., & Floreano, D. (2009). Evolved swarming without positioning information: An application in aerial communication relay. *Autonomous Robots*, 26(1), 21–32.
- Hebb, D. O. (1949). *The organization of behavior: A neuropsychological theory*. Chapman & Hall.
- Hoff, N., Wood, R., & Nagpal, R. (2013). Distributed colony-level algorithm switching for robot swarm foraging. In N. Correll, M. Schwager, & M. Otte (Eds.), *Distributed autonomous robotic systems* (pp. 417–430). Springer.
- Hölldobler, B., & Wilson, E. O. (1990). *The ants*. Harvard University Press.
- Hollnagel, E., & Woods, D. D. (2005). *Joint cognitive systems. Foundations of cognitive systems engineering*. CRC Press.
- Howard, A., Matarić, M. J., & Sukhatme, G. S. (2002). An incremental self-deployment algorithm for mobile sensor networks. *Autonomous Robots*, 13(2), 113–126.
- Hubel, D. H., & Wiesel, T. N. (1962). Receptive fields, binocular interaction and functional architecture in the cat's visual cortex. *The Journal of Physiology*, 160(1), 106–154.
- Hutchins, E. (1995). *Cognition in the wild*. MIT Press.
- Ijspeert, A. J. (2008). Central pattern generators for locomotion control in animals and robots: A review. *Neural Networks*, 21(4), 642–653.
- Ioannou, P. A., & Sun, J. (1995). *Robust adaptive control*. Prentice Hall.
- Johnson-Laird, P. N. (1983). *Mental models: Towards a cognitive science of language, inference, and consciousness*. Cambridge University Press.
- Kahneman, D. (2011). *Thinking, fast and slow*. Macmillan.
- Karavas, G. K., Larsson, D. T., & Artemiadis, P. (2017). A hybrid BMI for control of robotic swarms: Preliminary results. In *2017 IEEE/RSJ International Conference on Intelligent Robots and Systems (IROS)* (pp. 5065–5075). IEEE.
- Kieliba, P., Clode, D., Maimon-Mor R. O., & Makin, T. R. (2021). Robotic hand augmentation drives changes in neural body representation. *Science Robotics*, 6(54), eabd7935. doi:10.1126/scirobotics.abd7935
- Klein, G., Woods, D. D., Bradshaw, J. M., Hoffman, R. R., & Feltovich, P. J. (2004). Ten challenges for making

- automation a 'team player' in joint human-agent activity. *IEEE Intelligent Systems*, 19(6), 91–95.
- Kolling, A., Walker, P., Chakraborty, N., Sycara, K., & Lewis, M. (2015). Human interaction with robot swarms: A survey. *IEEE Transactions on Human-Machine Systems*, 46(1), 9–26.
- Kwakkel, G., Kollen, B. J., & Krebs, H. I. (2008). Effects of robot-assisted therapy on upper limb recovery after stroke: A systematic review. *Neurorehabilitation and Neural Repair*, 22(2), 111–121.
- Le Goc, M., Kim, L. H., Parsaei, A., Fekete, J.-D., Dragicevic, P., & Follmer, S. (2016). Zoids: Building blocks for swarm user interfaces. In *Proceedings of the 29th Annual Symposium on User Interface Software and Technology* (pp. 97–109). Association for Computing Machinery.
- Lebedev, M. A., & Nicolelis, M. A. L. (2017). Brain-machine interfaces: From basic science to neuroprostheses and neurorehabilitation. *Physiological Reviews*, 97, 767–837.
- Lee, S. K., Fekete, S. P., & McLurkin, J. (2016). Structured triangulation in multi-robot systems: Coverage, patrolling, voronoi partitions, and geodesic centers. *The International Journal of Robotics Research*, 35(10), 1234–1260.
- Lewis, M. (2013). Human interaction with multiple remote robots. *Reviews of Human Factors and Ergonomics*, 9(1), 131–174.
- Lewis, M., Sycara, K., & Walker, P. (2018). The role of trust in human-robot interaction. In J. Scholz, D. J. Reid, & H. A. Abbass (Eds.), *Foundations of trusted autonomy* (pp. 135–159). Springer.
- Licklider, J. C. R. (1960). Man-computer symbiosis. *IRE Transactions on Human Factors in Electronics, HFE-1*, 4–11.
- Lilienthal, A., & Duckett, T. (2004). Experimental analysis of gas-sensitive Braitenberg vehicles. *Advanced Robotics*, 18(8), 817–834.
- Maravita, A., & Iriki, A. (2004). Tools for the body (schema). *Trends in Cognitive Sciences*, 8(2), 79–86.
- Marder, E., & Bucher, D. (2001). Central pattern generators and the control of rhythmic movements. *Current Biology*, 11(23), R986–R996.
- Marr, D. (1982). *Vision: A computational investigation into the human representation and processing of visual information*. Freeman.
- Mathews, N., Christensen, A. L., O'Grady, R., Mondada, F., & Dorigo, M. (2017). Mergeable nervous systems for robots. *Nature Communications*, 8(1), 1–7.
- Miller, L. E., Montroni, L., Koun, E., Salemme, R., Hayward, V., & Farnè, A. (2018). Sensing with tools extends somatosensory processing beyond the body. *Nature*, 561(7722), 239–242.
- Minsky, M. (1988). *Society of mind*. Simon & Schuster.
- Molins, P., Stillman, N., & Hauert, S. (2019). Trail formation using large swarms of minimal robots. *Cybernetics and Systems*, 50(8), 693–710.
- Monaco, J. D., Hwang, G. M., Schultz, K. M., & Zhang, K. (2020). Cognitive swarming in complex environments with attractor dynamics and oscillatory computing. *Biological Cybernetics*, 114(2), 269–284.
- Mondada, L., Karim, M. E., & Mondada, F. (2016). Electroencephalography as implicit communication channel for proximal interaction between humans and robot swarms. *Swarm Intelligence*, 10(4), 247–265.
- Murphy, R. R. (2019). *Introduction to AI robotics*. MIT Press.
- Musić, S., & Hirche, S. (2017). Control sharing in human-robot team interaction. *Annual Reviews in Control*, 44, 342–354.
- Nagi, J., Ngo, H., Gambardella, L. M., & Di Caro, G. A. (2015). Wisdom of the swarm for cooperative decision-making in human-swarm interaction. *Proceedings – IEEE International Conference on Robotics and Automation*, 2015, 1802–1808.
- Nam, C., Walker, P., Li, H., Lewis, M., & Sycara, K. (2019). Models of trust in human control of swarms with varied levels of autonomy. *IEEE Transactions on Human-Machine Systems*, 50(3), 194–204.
- Neisser, U. (1976). *Cognition and reality*. W.H. Freeman.
- Nikolić, D. (2015). Practopoiesis: Or how life fosters a mind. *Journal of Theoretical Biology*, 373, 40–61.
- Nolfi, S., & Floreano, D. (2000). *Evolutionary robotics: The biology, intelligence, and technology of self-organizing machines*. MIT Press.
- Nouyan, S., Campo, A., & Dorigo, M. (2008). Path formation in a robot swarm. *Swarm Intelligence*, 2(1), 1–23.
- Oh, H., Shirazi, A. R., Sun, C., & Jin, Y. (2017). Bio-inspired self-organising multi-robot pattern formation: A review. *Robotics and Autonomous Systems*, 91, 83–100.
- Otte, M. (2018). An emergent group mind across a swarm of robots: Collective cognition and distributed sensing via a shared wireless neural network. *The International Journal of Robotics Research*, 37(9), 1017–1061.
- Parasuraman, R., Sheridan, T., & Wickens, C. D. (2000). A model for types and levels of human interaction with automation. *IEEE Transactions on Systems, Man, and Cybernetics – Part A: Systems and Humans*, 30(3), 286–297.
- Parker, L. E. (2007). Distributed intelligence: Overview of the field and its application in multi-robot systems. In *AAAI Fall Symposium: Regarding the Intelligence in Distributed Intelligent Systems* (pp. 1–6). AAAI Press.
- Passino, K. M., Seeley, T. D., & Visscher, P. K. (2008). Swarm cognition in honey bees. *Behavioral Ecology and Sociobiology*, 62(3), 401–414.
- Payton, D., Daily, M., Estowski, R., Howard, M., & Lee, C. (2001). Pheromone robotics. E. Şahin, & W. M. Spears (Eds.), *Autonomous Robots*, 11(3), 319–324.
- Payton, D., Estkowski, R., & Howard, M. (2004). Pheromone robotics and the logic of virtual pheromones. In E. Şahin, & W. M. Spears (Eds.), *International Workshop on Swarm Robotics* (pp. 45–57). Springer.
- Pearson, K., & Gordon, J. (2013). Locomotion. In E. Kandel, J. Schwartz, T. Jessell, S. Siegelbaum & A. Hudspeth (Eds.), *Principles of neural science* (5th ed., pp. 812–834). McGraw-Hill.
- Penders, J., Alboul, L., Witkowski, U., Naghsh, A., Saez-Pons, J., Herbrechtsmeier, S., & El-Habbal, M. (2011). A robot swarm assisting a human fire-fighter. *Advanced Robotics*, 25(1–2), 93–117.
- Pfeifer, R., & Bongard, J. (2006). *How the body shapes the way we think: A new view of intelligence*. MIT Press.
- Podevijn, G., O'Grady, R., Mathews, N., Gilles, A., Fantini-Hauwel, C., & Dorigo, M. (2016). Investigating the effect of increasing robot group sizes on the human psychophysiological state in the context of human-swarm interaction. *Swarm Intelligence*, 10(3), 193–210.

- Pourmehr, S., Monajjemi, V. M., Vaughan, R., & Mori, G. (2013). 'You two! Take off!': Creating, modifying and commanding groups of robots using face engagement and indirect speech in voice commands. In *IEEE International Conference on Intelligent Robots and Systems* (pp. 137–142). IEEE.
- Powers, W. T. (1973). *Behavior: The control of perception*. Aldine.
- Rasmussen, J., Pejtersen, A. M., & Goodstein, L. P. (1994). *Cognitive systems engineering*. Wiley.
- Reina, A., Bose, T., Trianni, V., & Marshall, J. A. R. (2018). Psychophysical laws and the superorganism. *Scientific Reports*, 8(1), 1–8.
- Reina, A., Cope, A. J., Nikolaidis, E., Marshall, J. A. R., & Sabo, C. (2017). Ark: Augmented reality for kilobots. *IEEE Robotics and Automation Letters*, 2(3), 1755–1761.
- Reina, A., Miletitch, R., Dorigo, M., & Trianni, V. (2015). A quantitative micro–macro link for collective decisions: The shortest path discovery/selection example. *Swarm Intelligence*, 9(2–3), 75–102.
- Richards, W. A. (2015). *Anigras: Experiments in cooperative cognitive architecture*. MIT Press.
- Rolls, E. T. (2016). *Cerebral cortex: Principles of operation*. Oxford University Press.
- Roundtree, K. A., Goodrich, M. A., & Adams, J. A. (2019). Transparency: Transitioning from human–machine systems to human–swarm systems. *Journal of Cognitive Engineering and Decision Making*, 13(3), 171–195.
- Rubenstein, M., Cornejo, A., & Nagpal, R. (2014). Programmable self-assembly in a thousand-robot swarm. *Science*, 345(6198), 795–799.
- Russell, S., & Norvig, P. (2016). *Artificial intelligence: A modern approach* (3rd ed.). Pearson.
- Şahin, E. (2004). Swarm robotics: From sources of inspiration to domains of application. In E. Şahin, & W. M. Spears (Eds.), *International workshop on swarm robotics* (pp. 10–20). Springer.
- Sanes, J. R., & Jessell, T. M. (2013a). Experience and the refinement of synaptic connections. In E. Kandel, J. Schwartz, T. Jessell, S. Siegelbaum, & A. Hudspeth (Eds.), *Principles of Neural Science* (5th ed., pp. 1259–1283). McGraw-Hill.
- Sanes, J. R., & Jessell, T. M. (2013b). The growth and guidance of axons. In E. Kandel, J. Schwartz, T. Jessell, S. Siegelbaum, & A. Hudspeth (Eds.), *Principles of neural science* (5th ed., pp. 1209–1232). McGraw-Hill.
- Sasaki, T., & Pratt, S. C. (2018). The psychology of superorganisms: Collective decision making by insect societies. *Annual Review of Entomology*, 63, 259–275.
- Schettler, A., Raja, V., & Anderson, M. L. (2019). The embodiment of objects: Review, analysis, and future directions. *Frontiers in Neuroscience*, 13, Article 1332.
- Schmickl, T., & Crailsheim, K. (2008). Trophallaxis within a robotic swarm: Bio-inspired communication among robots in a swarm. *Autonomous Robots*, 25(1–2), 171–188.
- Schranz, M., Umlauft, M., Sendel, M., & Elmenreich, W. (2020). Swarm robotic behaviors and current applications. *Frontiers in Robotics and AI*, 7, Article 36.
- Schwarz, J., & Fuchs, S. (2017). Multidimensional real-time assessment of user state and performance to trigger dynamic system adaptation. In D. Schmorow & C. Fidopiastis (Eds.), *Lecture notes in computer science (including subseries Lecture Notes in Artificial Intelligence and Lecture Notes in Bioinformatics)* (pp. 383–398). Springer.
- Senge, P. M. (2006). *The fifth discipline: The art and practice of the learning organization*. Currency.
- Seth, A. K. (2013). Interoceptive inference, emotion, and the embodied self. *Trends in Cognitive Sciences*, 17(11), 565–573.
- Seth, A. K. (2014). *The cybernetic Bayesian brain*. MIND Group.
- Sheridan, T. B. (1992). *Telerobotics, automation, and human supervisory control*. MIT Press.
- Simon, H. (1982). *The sciences of the artificial*. MIT Press.
- Slavkov, I., Carrillo-Zapata, D., Carranza, N., Diego, X., Jansson, F., Kaandorp, J., Hauert, S., & Sharpe, J. (2018). Morphogenesis in robot swarms. *Science Robotics*, 3(25), Article eaau9178.
- Sperati, V., Trianni, V., & Nolfi, S. (2011). Self-organised path formation in a swarm of robots. *Swarm Intelligence*, 5(2), 97–119.
- Stirling, T., Wischmann, S., & Floreano, D. (2010). Energy-efficient indoor search by swarms of simulated flying robots without global information. *Swarm Intelligence*, 4(2), 117–143.
- Sumpter, D. J. T. (2006). The principles of collective animal behaviour. *Philosophical Transactions of the Royal Society B: Biological Sciences*, 361(1465), 5–22.
- Szymanski, M., Breitling, T., Seyfried, J., & Wörn, H. (2006). Distributed shortest-path finding by a microrobot swarm. In M. Dorigo, L. M. Gambardella, M. Birattari, A. Martinioli, R. Poli, & T. Stützle (Eds.), *International Workshop on Ant Colony Optimization and Swarm Intelligence* (pp. 404–411). Springer.
- Tavakoli, A., Nalbandian, H., & Ayanian, N. (2016). Crowd-sourced coordination through online games. In *2016 11th ACM/IEEE International Conference on Human-Robot Interaction (HRI)* (pp. 527–528). IEEE.
- Trianni, V., Tuci, E., Passino, K. M., & Marshall, J. A. (2011). Swarm Cognition: An interdisciplinary approach to the study of self-organising biological collectives. *Swarm Intelligence*, 5(1), 3–18.
- Tsykunov, E., Agishev, R., Ibrahimov, R., Labazanova, L., Tleugazy, A., & Tsetserukou, D. (2019). Swarmitouch: Guiding a swarm of micro-quadrotors with impedance control using a wearable tactile interface. *IEEE Transactions on Haptics*, 12(3), 363–374.
- Valentini, G., Ferrante, E., & Dorigo, M. (2017). The best-of-n problem in robot swarms: Formalization, state of the art, and novel perspectives. *Frontiers in Robotics and AI*, 4, Article 9.
- Varela, F. G., Maturana, H. R., & Uribe, R. (1974). Autopoiesis: The organization of living systems, its characterization and a model. *Biosystems*, 5(4), 187–196.
- Verbruggen, M. (2019). *The question of swarms control: Challenges to ensuring human control over military swarms*. EU Non-Proliferation and Disarmament Consortium.
- Vicente, K. J., & Rasmussen, J. (1992). Ecological interface design: Theoretical foundations. *IEEE Transactions on Systems, Man, and Cybernetics*, 22(4), 589–606.
- Villani, V., Capelli, B., Secchi, C., Fantuzzi, C., & Sabatini, L. (2020). Humans interacting with multi-robot systems: A natural affect-based approach. *Autonomous Robots*, 44(3), 601–616.

- Walker, P., Nunnally, S., Lewis, M., Chakraborty, N., & Sycara, K. (2013). Levels of automation for human influence of robot swarms. *Proceedings of the Human Factors and Ergonomics Society Annual Meeting*, 57, 429–433.
- Walker, P., Nunnally, S., Lewis, M., Kolling, A., Chakraborty, N., & Sycara, K. (2012). Neglect benevolence in human control of swarms in the presence of latency. In *2012 IEEE International Conference on Systems, Man, and Cybernetics (SMC)* (pp. 3009–3014). IEEE.
- Werger, B. B., & Mataric, M. J. (1996). Robotic food chains: Externalization of state and program for minimal-agent foraging. *From Animals to Animats*, 4, 625–634.
- Wolpaw, J. R., Birbaumer, N., McFarland, D. J., Pfurtscheller, G., & Vaughan, T. M. (2002). Brain–computer interfaces for communication and control. *Clinical Neurophysiology*, 113(6), 767–791.
- Woods, D. D. (2015). Four concepts for resilience and the implications for the future of resilience engineering. *Reliability Engineering & System Safety*, 141, 5–9.
- Woods, D. D. (2018). The theory of graceful extensibility: Basic rules that govern adaptive systems. *Environment Systems and Decisions*, 38(4), 433–457.
- Woods, D. D., & Hollnagel, E. (2006). *Joint cognitive systems: Patterns in cognitive systems engineering*. CRC Press.
- Yang, G. Z., Bellingham, J., Dupont, P. E., Fischer, P., Floridi, L., Full, R., Jacobstein, N., Kumar, V., McNutt, M., Merrifield, R., Nelson, B. J., Scassellati, B., Taddeo, M., Taylor, R., Veloso, M., Wang, Z. L., & Wood, R. (2018). The grand challenges of science robotics. *Science Robotics*, 3(14), Article eaar7650.
- Zhu, W., Allwright, M., Heinrich, M. K., Oğuz, S., Christensen, A. L., & Dorigo, M. (2020). Formation control of UAVs and mobile robots using self-organized communication topologies. In M. Dorigo, T. Stützle, M. J. Blesa, Ü. Blum, E. Hamann, M. K. Heinrich, & O. Strobel (Eds.), *International Conference on Swarm Intelligence* (pp. 306–314). Springer.

## About the Authors



**Jonas D Hasbach** is a research scientist with the department of Human-Machine-Systems at the Fraunhofer Institute for Communication, Information Processing and Ergonomics FKIE, Wachtberg, Germany. Prior to his employment at the Fraunhofer FKIE, he received his MSc in Cognitive Neuroscience from Maastricht University, the Netherlands in 2016. His main research topics are the analysis and design of complex command and control systems and human-autonomy teaming. In addition, he pursues a PhD at the department of Computer Science VI, University of Bonn, Germany where he focuses on bio-inspired human-swarm interaction.



**Maren Bennewitz** is a professor of Computer Science at the University of Bonn, Germany, and head of the Humanoid Robots Lab. She got her PhD in Computer Science from the University of Freiburg in 2004. Before she moved to Bonn in 2014, she was a Postdoc and assistant professor at the University of Freiburg. The focus of her research lies on robots acting in human environments. In the last few years, she has been developing several innovative solutions for robotic systems co-existing and interacting with humans. Among them are techniques for efficient navigation with humanoid and wheeled robots as well as for reliably detecting and tracking humans from sensor data and analysing their motions. She has published over 100 peer-reviewed scientific papers. She regularly serves as editor or associate editor for top-tier robotics conferences, in the editorial board of top-tier robotics journals, and as programme committee member for international workshops. She also serves as reviewer for the European Commission (EC) and the German Research Foundation (DFG) and has been principal investigator (PI) in several national and European projects.