The perceived effects of augmented trail sensing and mood recognition abilities in a human–fish biohybrid system

Xin Hu, Jinxin Yang, Zhihua Song, Qian Wang, Ziyue Chu, Lei Zhang, Daoyuan Lin, Yangyang Xu, Longfei Liang and Wen-Chi Yang

Intelligent Systems Lab, NeuHelium Co., Ltd. Shanghai 200090 People's Republic of China

* Author to whom any correspondence should be addressed.

E-mail: wenchi.yang@neuhelium.com

Keywords: artificial life, biohybrid systems, perception enhancement, cross-species interaction

Abstract

The use of technologies to enhance human and animal perception has been explored in pioneering research about artificial life and biohybrid systems. These attempts have revealed that augmented sensing abilities can emerge from new interactions between individuals within or across species. Nevertheless, the diverse effects of different augmented capabilities have been less examined and compared. In this work, we built a human–fish biohybrid system that enhanced the vision of the ornamental fish by projecting human participants onto the arena background. In contrast, human participants were equipped with a mixed-reality device, which visualized individual fish trails (representing situation-oriented perceptions) and emotions (representing communication-oriented perceptions). We investigated the impacts of the two enhanced perceptions on the human side and documented the perceived effects from three aspects. First, both augmented perceptions considerably increase participants' attention toward ornamental fish, and the impact of emotion recognition is more potent than trail sense. Secondly, the frequency of human–fish interactions increases with the equipped perceptions. The mood recognition ability on the human side can indirectly promote the recorded positive mood of fish. Thirdly, most participants mentioned that they felt closer to those fish which had mood recognition ability, even if we added some mistakes in the accuracy of mood recognition. In contrast, the addition of trail sensing ability does not lead to a similar effect on the mental bond. These findings reveal several aspects of different perceived effects between the enhancements of communication-oriented and situation-oriented perceptions.

1. Introduction

Natural organisms (animals, plants, microorganisms) have evolved with diverse sensing capabilities to interact with conspecifics and respond to the environment (Evans 1996, Bijlsma and Loeschcke 2013, Richardson et al 2014). These inherent perceptions are vital to organisms in nature because they directly determine the channels and amounts of information that can be acquired from the surroundings. Naturally, these perceptions may deteriorate with time due to aging or impairment (Roberts and Allen 2016). However, based on technology, enhancement of perceptions becomes possible such that natural organisms can respond to their environment in a new way. Various studies (Macrae and Bodenhausen 2001, Corbetta and Shulman 2002, Bodenhausen and Hugenberg 2009, Rensink 2013, Stephenson et al 2021, Wolfe et al 2022) have shown that enhanced perceptions can concentrate human attention and change people's understanding of their social environment, but different perceptions have various effects on these changes. This article aims to investigate and compare how two distinct types of enhanced perception (one towards situational awareness and the other towards social communication) affect the cognition and behavior of humans when they interact with a tank of ornamental fish.

1.1. Background to perception

Perception is a natural organism's ability to organize, identify and interpret presented information gathered from its sensors for understanding its living surroundings. This ability is essential to most natural
organisms, as it can directly determine cognition and behavior (Bruner and Postman 1949). For human beings, it is crucial to perceive detailed information about the behavioral semantics and emotional states of others for the achievement of self-regulation (Horne 2012). For other animals, such as elephants (e.g. smell to pick up unusual odors and avoid danger; Bates et al 2007) and rats (e.g. touch with their whiskers to find food and to communicate with conspecifics; Smith and Alloway 2013), perceptions are vital for survival and reproduction. Each creature has particular types of perceptions, by which a (limited) part of local information is extracted for the understanding of circumstances. Taking tactile perception for example, mimusas (Cabill Jr et al 2013) and snails (Logunov and Konnov 1983) can only sense the strength of pressure, ants (Ratnieks 2007) and bees (Esch et al 2001) have evolved to communicate through tactile perception, while gorillas (Clark et al 2019) and dogs (Byrne et al 2017) can identify intention and emotion through touch stimuli conducted from haptic cells in their skin.

In general, we can divide perceptions into situation-oriented and communication-oriented functions by the utility of the perceived information. Situation-oriented perception (situational awareness) refers to those functions concerning awareness of information about the surroundings (Bischoff and Graefe 1999). Trial sensing is one example of integrated situation-oriented perception that is well developed in predatory mammals such as bears, wolves and dogs (Standing et al 1970) and various species of ant. Communication-oriented perception (social perception) refers to the perceptions built up to achieve interaction, communication and social cognition between individuals of the same species (Mackie et al 2000), such as mood recognition and language acquisition. The retrieved semantics of these perceptions vary among individuals, because no uniform objective evaluation or measurement can be applied even if they occur under the same circumstances. For example, past working memory may have an influence on individuals’ emotions when in a group of people in the same physical environment because the received perception stimuli may be augmented or neglected when they pass through the memory part of the brain (Kessel et al 2016). People’s emotions may also be influenced by other people in the same environment, and interaction among individuals can reinforce group-level emotions (Smith and Mackie 2016).

1.2. Perception enhancement
From the perspective of individuals, specific sensory impairment, usually due to injury or aging, is common in nature (Roberts and Allen 2016), and the harm caused by the loss of certain perceptions can differ from species to species. For example, deprivation of vision is fatal to shoal fish because vision helps shoal fish maintain coordinated movement, while the loss of visual ability can be less critical to ants since ants can still survive with tactile sensing (Seidl and Wegner 2006, Clifton et al 2020). Nevertheless, although some loss of certain perceptions can be adapted to afterwards, it can still change individual cognition and behaviors to some extent. A decline in taste perception in females can result in a change in eating habits (Pepino et al 2014) and older domesticated cats often suffer from feline cognitive dysfunction, which causes disturbances in sleeping patterns and reduced activity due to decline in sight and hearing perception (Chen et al 2015). Even for individuals with normal-level perceptions, enhanced perceptions can be advantageous in biological evolution, as these perceptions can provide individuals with more information enabling them to warn about and understand nearby and neglected dangers. For human beings, perception enhancement can not only help people with perception impairment lead a normal life but can also lead to a deeper awareness and understanding of the surrounding environment (Tyan et al 2014, Chu et al 2018).

The importance of perception enhancement to normal people lies in a better sense of the environment, which changes people’s cognition and behavior compared with normal levels of perception. In psychology, cognition is categorized into attention, social cognition, memory, executive function and psychomotor speed according to its diverse functions (Kihlstrom 1987, Adolphs 2006, Bodenhausen and Hugenberg 2009). Attention is the process of allocating limited ‘resources’ to select and track a particular item based on perceived global importance, associated with the intensity of perception (Rensink 2013). Irrelevant information can be noise to perception, and enhancement of relevant perception can increase the attention level to the target. For object detection tasks, individuals can detect the target more precisely when provided with advanced feature information (such as cueing) about the target location (Corbetta and Shulman 2002, Rensink 2013). The biological essence behind attentional shift is eye movement guided by peripheral information from enhanced perception (Wolfe et al 2022). On the other hand, social cognition is the observer’s ability to detect the state of others, such as emotion, traits or thoughts (Bodenhausen and Hugenberg 2009). Perceptions play an essential role in bridging between outside social targets and psychologically meaningful representations of inner experience, which directly determine social cognition (Bodenhausen and Hugenberg 2009). Research has also shown that category perception can simplify human understanding of the complex social world (Macrae and Bodenhausen 2001). In this article, we aim to investigate how human attention and social
cognition can be changed by different enhanced perceptions.

To achieve multi-perception enhancement beyond nature, different high-tech devices with mixed-reality (MR) technology have been applied in various domains. In medical practice, a surgery simulation system has been developed with augmented reality (AR) and HoloLens to enhance surgeons’ visual and audio perceptions, mixing virtual surgery scenarios with real actions (Condino et al. 2018). In the field of engineering, to obtain a better understanding of stiffness for tangible objects, a MR system renders stiffness with a two degree of freedom wearable tactile display for the finger (De Tinguy et al. 2018). In the field of education, MR/virtual reality (VR) devices are used in classes to increase students’ attention levels, which enhances their learning experience and engages them in active learning Azhar et al. (2018).

1.3. Biohybrid systems

A biohybrid system is a system with biology–machine interaction containing both biological and non-biological components. The concept comes from the idea of achieving collaboration between artificial systems and group-living animals by perceiving, communicating and interacting with animals (Halloy et al. 2013). Biohybrid systems help us understand how a biological apparatus (e.g. muscles) can work with the interaction of integrated technologies. Biological behavioral characteristics can be revealed in biohybrid systems. Furthermore, in an artificial system robots can also be used to manipulate the behavior of groups of living organisms to accomplish particular tasks (Halloy et al. 2013, Romano et al. 2019).

Two current topics in biohybrid system research are ‘biohybrid organisms’ and ‘animal–robot mixed societies’. ‘Biohybrid organisms’ focuses on integrating artificial devices with individual living organisms and ‘animal–robot mixed societies’ focuses on adding artificial devices (usually biomimetic robots) to biological communities (Romano et al. 2019). For biohybrid organisms, an integrated sensor array has been applied to enhance individuals’ olfactory and tactile perception capability (Liu et al. 2012, Lucarotti et al. 2013). Other researchers have studied rodents (Nickell et al. 2007, Guo et al. 2016, Zhang et al. 2018) and substitutes for real skin (Cheneler et al. 2014, Low et al. 2019, Jang et al. 2020). However, this research has mainly focussed on improving the method by which such enhancements can be achieved and investigating the potential electrophysiological mechanisms, rather than studying the impact of the enhanced perceptions. For animal–robot mixed societies, biomimetic robotic bees that can imitate the communication of bees (the waggle dance) (Michelsen et al 1992, Landgraf et al 2011, 2012, Gripparić et al 2017, Lazic and Schmickl 2021) have been introduced into the hive to avoid contact with polluted nectar sources and enhance the bee colony’s perception of local environments. This builds up an artificial ecological system, such as HIVEPOLIS (Ilgün et al 2021). Other attempts to promote human–fish interaction have been made by controlling the position and number of bubbles in a fish tank according to observers’ actions (‘Bubble Talk’) (Ko et al. 2018) and by displaying the inferred emotions of fish to enhance observers’ emotional perception of the fish (‘AffectiveNemo’) (Isokawa et al. 2019).

2. Experiments

2.1. Subjects

This research recruited 34 participants (21 men and 13 women aged 22–44 years; mean age 29 years, standard deviation ±6 years). Twenty of the participants had pets, whereas 14 had never owned one. Twenty-five individuals had prior familiarity with VR or AR, while the remaining nine had no prior experience.

2.2. Apparatus

A square interaction platform (9 m\(^2\)) was set on a 1.25 m high table, as shown in figure 1(a). Five goldfish (Carassius auratus) with distinct appearances were placed in a glass tank (0.9 m wide × 0.45 m deep × 0.45 m high) with water between 15°C and 20°C, depicted in figure 1(d). An oxygen pump and an aquarium filter were installed to provide the fish with a suitable living environment (figure 1(b)). A LCD display (MI L43M5-EK) was set up behind the aquarium to give the fish an interaction stimulus.

Sensor arrays, including cameras, depth cameras and computers, were deployed in the environment, mounted either above or around the glass tank. As depicted in figure 2, these sensors were utilized to track the movements of humans and fish, analyze sensory data and enhance perceptions in real time. An ordinary camera (aoniA30 HD 1080P) feeds data about humans to the computer at the edge of the tank. The depth camera (ZED 2) is mounted above the tank to gather information about the depth of the fish while simultaneously collecting live photographs of the tank. A computer (Intel(R) Core(TM) i7-9700K CPU@3.60 GHz, NVIDIA GeForce RTX 3080) used the artificial intelligence (AI) algorithm Yolo V5 to determine the location of the fish based on the depth information provided by the depth camera and to identify the position of participants wearing reflective vests with the regular camera.

A MR device (HoloLens 2) was used to blend the processed and inferred perceptual information with the real world and immersively enhance the participants’ perceptions by allowing them to immediately view the new perceptual information in the real world. HoloLens 2 included a gaze tracker (60 FPS)
Figure 1. A biohybrid system builds a bridge to promote interactions between humans and fish in which artificial intelligence analyzes the information obtained by sensor arrays and feedback to enhance the perception of humans and fish. (a) Side view of the biohybrid environment; (b) top view of the biohybrid environment; (c) biohybrid environment in use; (d) common goldfish (*Carassius auratus*) in our biohybrid system.

Figure 2. The biohybrid system constructed a bridge to facilitate interactions between people and fish, in which artificial intelligence (AI) evaluates the information gathered by a sensor array and feedback to improve human and fish perception. The sensor array captured human and fish movement data and sent them to computers. The algorithm Yolo V5 was utilized to determine the locomotion of fish and inferred the physiological states of the fish. A mixed reality device, HoloLens 2, was used to enhance human situation-oriented and communication-oriented perceptions. The integrated gaze tracker helped monitor human attention during various settings in the experiments.

that facilitated the generation of stimuli and analysis of human physical attention.

2.3. Stimuli
There were two types of stimuli in the experiment, situation-oriented and communication-oriented, both generated by HoloLens 2. To aid in detecting the subject of human attention, each fish was wrapped in a cube (0.0675 m³). The gaze tracker of HoloLens 2 follows the fixations of participants, and HoloLens 2 displays matching stimuli when the fixations are inside the rectangle.
The situation-oriented stimulus was a retracing trail that displayed the position and physical activity of fish in the past. As shown in figure 3, the different times for which the trail was retraced were utilized to control the levels of situation-oriented stimulus: (a) 2 s back (RT1), (b) 4 s back (RT2) and (c) 8 s back (RT3). Similar to the fish, the retracing trail comprises numerous cubes \( (8.4375 \times 10^{-4} \text{ m}^3) \) to help identify attention. The retracing trail was shown immediately when fixations were on the fish and erased when fixations moved away.

The communication-oriented stimulus was a mood tag (shown as emojis) developed based on the behavior of the fish. As shown in figure 4(a), five emojis indicating neophobic, aroused, freezing, fright and normal were used to express the physiological states of fish to humans (Laming and Savage 1980, Kim et al 2014). Randomness was implemented to regulate varying levels of expression. Emojis were selected randomly at three levels of expression accuracy: (a) randomly selecting one emoji out of four emojis that do not contain the predicted one based on the physiological state of the fish (MT1), in which each of the four emoji has a 25% chance of being selected; (b) selecting one emoji out of five with proportional probability (MT2), in which the predicted emoji has a 50% chance of being chosen while the last four have a 12.5% chance of being chosen individually; and (c) displaying the predicted emoji (MT3). The mood tag is likewise encompassed by a globe \((\text{radius } 0.0108 \text{ m})\) to aid in the placement of the individual’s gaze. Similar to the situation-oriented stimulus, the mood tag was quickly shown while fixations were on the fish and disappeared when fixations were no longer on the fish.

In addition to enhancing human perceptions, fish were also exposed to stimuli to boost the efficiency of human–fish encounters and to determine the responses of the fish to human contact. Previous studies have demonstrated that light may alter fish behavior (Popper and Carlson 1998, Romano and Stefanini 2022b) and that fish can receive social and emotional support from robotic fish (Romano and Stefanini 2022a). To guarantee the validity of the interaction, we designed a means for people to interact with fish. The designed interaction included both stimuli above. The monitor’s backdrop color featured two modes, dark and light. At the beginning of the trial, three artificial fish of various hues swam in response to human movement on a black backdrop. The fake fish were there until the end of the trial. As soon as the user lifted his or her hand, the computer identified the gesture taken by the ordinary camera and changed the monitor’s backdrop between dark and light.

2.4. Procedure
Participants wearing a reflective vest and HoloLens 2 were instructed to freely examine the fish tank and engage with the fish while standing in front of the
participants' attention to objects in the aquarium, identifying human gaze or eye movement serves cognitively with fish. Four different combinations of situation-oriented and communication-oriented stimuli could be presented: (a) a single situation-oriented stimulus, Scene Trail, a retracing trail; (b) a communication-oriented stimulus, Scene Mood, a mood tag presented alone; (c) a situation-oriented stimulus and a communication-oriented stimulus, Scene M&T, in which both the backtracking trail and mood tag are shown; (d) Scene Null, no stimulus but the real world. In the cross-modal condition, there are nine combinations of stimulus levels. In both Scene Trail and Scene Mood, participants were exposed to three 60 s periods of the matching stimulus. In Scene M&T, three out of nine possible combinations were selected randomly for investigation so that participants would still experience three 60 s sessions. In Scene Null, participants experienced three 60 s periods. The observations in each scene are made in succession, uninterrupted. After each scene, participants were queried with a subjective survey (appendix A). To counteract any possible behavioral or psychological bias associated with time, four scenarios and three sessions in Scene Trail, Mood, and M&T were randomly ranked. The whole process for each person took around 25 min.

After each experiment, five types of information were gathered by the sensor arrays: (a) the gaze data from the gaze tracking of HoloLens 2, which contained participants’ fixations during the experiment (in experiments in which the mood tag or the trace trail was not displayed, the participants’ gaze fixations on the place where the mood tag or the trace trail was supposed to appear were taken into account); (b) the fish locomotion, which recorded the movement of fish by the depth camera; (c) the relative position of humans in front of the tank by the ordinary camera; (d) the time and movement of human engagement with fish recorded by the ordinary camera; and (e) the subjective surveys following each scene.

3. Results

3.1. Enhanced perceptions allow humans to engage cognitively with fish

3.1.1. Physical attention varies with different enhanced perceptions

Identifying human gaze or eye movement serves the ultimate objective of determining an individual’s attention. HoloLens 2 gaze data assisted in monitoring participants’ attention to objects in the aquarium, such as the fish encased in a rectangle, mood tags encased in a rectangle and retracing trails comprising many cubes. In 60 s, the number of fixations inside matched targets was compared to indicate the attention level of participants.

There are aspects that may influence attention, such as strong stimuli, movements, emotional strain, etc (Knudsen 2007). Our study suggests that situation-oriented perception may affect attention differently than communication-oriented perception. The retracing traces brought more attention to the fish itself, but the mood tag, regardless of whether the phrase was correct or not, drew more attention to the mood of the fish rather than the fish itself. Consequently, when two perceptual improvements were present, the focus was divided between the fish itself and the mood tag, which gave a distinct outcome. The intensification of perception might increase overall attention in superposition.

As indicated in figure 5, enhancing either situation-oriented stimuli (RT1, RT2, RT3) or communication-oriented stimuli (MT1, MT2, MT3) significantly increased participants’ attention (p-values are shown in tables B1(a) and (b)). However, the intensity of each stimulus is not positively correlated with attention, which means that a trail traced farther back or more accurate mood expression does not contribute to an increase in total attention. Regarding the situation-based stimulus, the retracing track received twice as much attention when it was traced 4 s earlier (RT2) as in Scene Null. The trail retraced 8 s back (RT3) seemed to redirect participants’ attention away from the fish and fish-related objects, resulting in a decrease in overall attention relative to RT1 and RT2. One potential explanation for this phenomenon is that the retracing of tracks sends varying amounts of information to which humans can attend. However, attention will be diverted when the quantity of information is much less or greater than expected.

The impact of mood tags on human attention was significantly greater than that of Scene Null (p-values shown in table B1(b)), in contrast to the effect of retracing tracks on human attention. The attention to trails has been greatly diverted to the attention to mood tags, whilst the attention to fish remains almost equal to Scene Null. Similar to the correlation found in Scene Trail, the intensification of the accuracy of mood expression does not have a positive correlation with attention either (MT1, MT2, MT3). Compared with Scene Null, humans place significantly more emphasis on the mood tags themselves, independent of the veracity of the expression.

The superposition of retracing trails and mood tags had distinct impacts compared with when just one of the perceptual enhancements is present. Compared with the session with only retracing trails (RT1,
Figure 5. The average physical attention paid to fish, the retraceing trail and the mood tag in various scenarios and sessions. Enhancing either the retraceing trail (RT1, RT2, RT3) or the mood tag (MT1, MT2, MT3) can considerably improve human physical attention compared with Scene Null. Retraceing trails (RT1, RT2, RT3) improves human focus on the fish themselves, while showing mood tags (MT1, MT2, MT3) enhances participants’ focus on the mood tags. Shorter retraceing trails (RT1, RT2) attract more attention than longer ones (RT3). The medium accuracy of mood expressions (MT2) generates more interest in mood tags than the other two (MT1, MT3). Compared with the sessions with only retraceing trails (RT1, RT3), adding more accurate mood expressions (RT1&MT1, RT1&MT3, RT3&MT1, RT3&MT3) greatly increased the overall attention. On the other hand, the addition of longer retraceing trails (RT1&MT1, RT1&MT3, RT3&MT1, RT3&MT3) does not show a significant increase in overall attention compared with the sessions with only mood tags (MT1, MT3).

In more detail, given the same retraceing trail stimulus, the addition of a mood tag extracts the attention to the fish and acts on the attention to the mood of the fish. On the other hand, adding retraceing trails (MT1 vs RT1&MT1, MT1 vs RT3&MT1, MT3 vs RT1&MT3, MT3 vs RT3&MT3) distracts human attention from the mood. Intensifying the situation-oriented stimulus helps increase human attention to the fish, while the increase is not as significant as intensifying the communication-oriented stimulus (p-values shown in tables B1(c)–(f)).

Therefore, situation-oriented perception affects human attention to the item itself, but communication-oriented perception affects human attention to the communication's semantics. The superposition of both senses does not equal the total effect of adding either a situation-oriented stimulus or a communication-oriented stimulus. The attention will oscillate between the object and the communication semantics, resulting in a general decline in attention. Intensifying communication-oriented perception is more effective than intensifying situation-oriented perception for boosting attention to objects and object-related things. In the superposition condition, attention is less than in the single perception condition.

3.1.2. Subjective attention varies with different enhanced perceptions

As well as measurable physical attention, there are internal factors in individuals which affect their attention, such as interests, the effort required by the task, trains of thought, etc (Knudsen 2002). The second question relating to attention (appendix A) is queried after each scene in subjective surveys. The scoring range for the question is from 1 to 5 (strongly disagree to strongly agree).

Since the average score for the attention-related question in each survey (appendix A), Scene Null (average 3.97), Scene Trail (average 3.76), Scene Mood (average 4.0) and Scene Mood&Trail (average 3.88), failed the Shapiro–Wilk test, the Friedman test was used to compare the four groups. According to table B3(a) in the appendix, these four groups have a common distribution. A Friedman post hoc Wilcoxon test was used to compare each pair of groups. The subjective attention to fish in Scene Mood is higher than in the other three. The inclusion of mood tags and trails led participants to believe that they paid less attention to the fish than to Scene Null. In Scene Mood&Trail, participants paid somewhat more attention to the fish than they did in Scene
Figure 6. Distributions of preset and free interaction occurring in four different scenes: when both perceptions were enhanced, participants tend to interact more frequently regardless of the methods of interaction. (a) The number of interaction behaviors that are predetermined in four scenes; (b) the number of interaction behaviors that are free in four scenes.

Trail, which differs from our description of physical attention above. Individuals' concern for communication semantics, i.e. mood, may account for their heightened awareness of fish rather than the retracing trails. Participants might treat the retracing trails as objects rather than part of the fish. This also explains why fish were believed to be the least attentive in Scene Trail.

In conclusion, focusing more on communication semantics is likely to communicate humanity and provide the impression that one is paying greater attention to the object. However, this impact is insufficient to counteract the mental distraction caused by other objects.

3.1.3. The subjective effect of social cognition before interaction with fish
In the studies, five emojis represented the physiological states of fish, which may also be inferred from the fish trails. As attention fluctuated between fish and fish-related objects, individuals were more inclined to engage with fish. In this research, the willingness to interact with fish, which reflects social cognition of humans, is also considered. The third question in each survey (A.2.3, A.3.3, A.4.3, A.5.3 in appendix A) inquired about participants' readiness to interact with these goldfish. The scoring range is from 1 to 5 (strongly disagree to strongly agree). The four groups failed the Shapiro–Wilk test and showed significant difference after applying the Friedman test (p-values shown in table B3(b)). According to the Friedman post hoc Wilcoxon test results in table B3(b), participants' propensity to engage with fish without mood tags and trails is considerably greater than that of the other three groups. Even without significance, there was a difference between Scene Trail and Scene Mood, indicating that individuals are more likely to interact with fish in scenes including both stimulus and mood tags than in settings containing simple trails. Since only one question was used to analyze participants' social cognition during the experiment, the result might be sensitive to the samples and individuals. More objective discussions about human–fish interactions are given in the following section.

3.2. Human–fish interaction
In human civilization, interaction is regarded as a dynamic, ever-changing series of social behaviors between individuals or groups. Likewise, for this biohybrid system, the interaction between humans and fish should be bidirectional, including not just human behavior to attract the attention of fish but also the changes in the physiological condition of the fish. As described in the preceding section, people may engage with fish using a specific gesture or they can do anything they choose. During the experiment, the interactional behaviors of the participants were counted to demonstrate expression from humans to fish. On the other hand, physiological variables were recorded and extrapolated to demonstrate the responsiveness of fish to human behavior. In addition to the data acquired from sensor arrays, subjective questionnaires indicated changes in the mental states of participants, allowing us to completely examine the mental reactions of people to various perception enhancements.

3.2.1. Interaction behavior from humans to fish
Participants were informed that they could act freely during the experiment, either interact or not, either use specific gestures to interact or use any other method to interact. Recorded data were analyzed by a third party to determine the number of times participants used other methods to interact with fish.

With interaction through specific gestures, a kernel density estimation (KDE) of the distribution of participants with the number of such interactions in each experiment was plotted in figure 6(a). It can be seen that most participants tend to interact through this method relatively few times (less than ten times) in an experiment, regardless of the scene. Two other noticeable points are that participants were
more likely to interact in a more frequent manner (more than 20 times) in scenes with both perceptions enhanced and were more likely to interact in a less frequent manner (less than 10 times) in scenes where only mood perception was enhanced. However, according to the results of the significance test in Table B2(a), the previous two points do not possess strong significance.

With other interaction methods that were freely used by participants, another KDE of the distribution of participants with the number of such interactions in each experiment was plotted in figure 6(b). It can be seen that this figure shows many similarities to figure 6(a). Most of participants tended to interact through this method relatively few times (less than 20 times) in an experiment, regardless of the scene. Participants were more likely to interact in a more frequent manner (more than 40 times) in scenes with both perceptions enhanced and were more likely to interact in a less frequent manner (less than 20 times) in the other scenes, according to the significance test results in table B2(b).

Such inconsistency of significance in different interactions, yet consistency in character, implies that enhancement of different perceptions may have an effect on participants’ behavior. However, for reasons which are not clear, such an effect did not cause a significant difference for every interaction method. This may be a confounding effect with multiple reasons, such as the experimental time not being long enough for participants to sufficiently demonstrate a strong difference for different perception enhancement or the change of cognition, proved by attention changes, may not always be reflected instantaneously or substantially.

3.2.2. Physiological states from fish to humans

During each experiment, which can be viewed as an interaction period, the change in the mood of fish during this interaction period was recorded and the mood response analyzed. Figure 7 shows the difference in mood response between different scenes. It is obvious that when comparing the differences between Scene Trail or Scene Mood and Scene Null (Trail vs Null and Mood vs Null), they share a pattern similarity in having more freezing and neophobic corresponding to less normal and fright as the mood response. Also, when comparing the difference between Scene Mood&Trail and Scene Trail or Scene Mood (M&T vs Trail and M&T vs Mood), they also share a pattern similarity, with Scene Mood&Trail have much less freezing than other moods in the mood response.

Such distinct synchronization in pattern may imply that participants tend to interrupt fish or cause neophobic behavior when they are in normal or fright mood, as freezing ‘is considered as a pause in ongoing behavior’ (Goodman and Weinberger 1973, Lamming and Savage 1980), when only one perception was enhanced. When both perceptions were enhanced, participants tend to interrupt less in exchange for causing other moods. However, as this difference
itself is reflected to some small extent, further research is needed.

3.2.3. The subjective feeling from humans
The fifth, sixth and seventh questions in each survey (A.2.5, A.3.5, A.4.5, A.5.5; A.2.6, A.3.6, A.4.6, A.5.6; A.3.7, A.4.7 in appendix A) investigated related aspects of participants' subjective opinions on interaction, with each question corresponding to participants' subjective opinion on the reaction of fish to participants' interaction action; participants' beliefs about understanding fish mood; and participants' desire to obtaining corresponding perception enhancement capability, respectively.

In question 5 participants were asked if they agree that a fish is reacting to their interaction. The static analysis showed that when both perceptions are enhanced, participants were more likely to believe that fish reacted to their interaction action compared with Scene Null; when enhancing trail perception, participants were less likely to believe so; when enhancing mood perception, participants' cognition shown no significant difference compared with Scene Null, according to the significance test results in table B3(d).

In question 6 participants were asked whether they believe they understand fish mood. The scene with enhanced mood perception significantly increased participants' confidence in their beliefs. Scenes with the only difference being whether trail perception was enhanced showed no significant effect on participants' confidence, according to the significance test in table B3(e). This suggests that the existence of communication-oriented perception enhancement, whether it is correct or not, can directly produce an effect of psychological suggestion on people's cognition, as the mood information presented to the participants contained some deliberately wrong information in some scenes. On the other hand, the augmentation of situation-oriented perception, trail perception, does not show such an effect, even though the mood of fish can be inferred from the fish trails. These results suggested that although both perceptions were enhanced through vision, their difference in characteristics, situation-oriented and communication-oriented, may cause this perception information to be processed by different parts of the brain, thus leading to a different effect on cognition.

Furthermore, evaluation of participants' perception preferences through question 7 shows that the perception that draws more attention, mood perception, is much preferred by participants, according to the significance test in table B3(f). Not only is this consistent with the effect of these perceptions on participants' attention levels, but it also combines the findings of questions 5 and 6 and suggests that participants prefer perceptions that enhance their social and communicative abilities. This may imply that social cognition is more important to participants than context-related cognition at a subconscious level.

4. Conclusion and discussion
This research explored the impacts of two distinct enhanced perceptions on people's attention and social cognition, respectively, which are found to result in a high rate of human–fish interaction with the support of MR devices and AI algorithm experiments. These findings reveal that perception enhancement is an artificial way to construct biological mental bonds between natural organisms and realize a biohybrid ecosystem.

The results in section 3.1 suggest that perception enhancement (both situation- and communication-oriented perception enhancement) increases human attention with different focusses in the experiment. Situation-oriented enhancement focuses human attention on relevant target items, while communication-oriented enhancement increases the communication's semantics. However, the superposition of both perception enhancements is a non-linear function of the input of each individual perception stimulus to attention. A general decline in attention occurs when both situation and communication are enhanced, which distracts participants' attention between the object and the communication semantics. Experimental statistics implies that enhancing communication-oriented perception is more effective than situation-oriented perception for boosting attention to target objects. Similar results were also reflected in the subjective questionnaires, with most participants agreeing that providing communication-oriented perception can help them concentrate more and have a greater interest in understanding the fish. This is because the presence of communication-oriented perception (i.e. mood tag) enhancement, whether it is correct or not, can directly produce a psychological suggestion-like effect on people's understanding of fish and thus concentrates their attention. Besides augmentation of attention, participants have also reported a raised willingness to interact with fish when provided with both stimulus and mood tags compared with only trail information in the subjective questionnaire. In the experiments in section 3.2, the increased willingness to interact triggers human–fish interaction. On one hand, participants are recorded to interact more frequently with fish when provided with both mood tag and trail. On the other hand, negative moods (i.e. freezing and neophobic) detected in the fish sharply decreased when interacting with human participants who were provided with both mood tags and trail information.

This research is a trial of applying advanced devices in biosystems to augment organisms' perception with the support of Artificial Life technologies (such as MR and AI algorithms), which can
be a benchmark for building a new ecosystem-level biohybrid system. For instance, this research can be used to enhance mutual understanding between humans and fish, thus leading to the formation of mental bonds among species and creating a new concept of the metaverse. The traditional metaverse aims to build a virtual human society with enhanced human perceptions, while this research is dedicated to creating a virtual society that promotes interaction and mental bonds between natural organisms and humans via enhanced cross-species perceptions; this could be a blueprint for a cross-species metaverse. Furthermore, such a cross-species metaverse will arouse human interest and increase humanity’s understanding of other species. With the mental bonds created, humans can better guide the behavior of endangered organisms in escaping environmental disasters or accomplishing tasks to protect local ecology and improve the natural environment in future biohybrid research. Apart from these applications, this paper also provides insights to measure human attention and breaks down human perception into situation-oriented and communication-oriented perception, so future discussion on human perception can be separated into the above two classes. The method in this paper for building a human–fish biohybrid system takes advantage of enhancement of visual perception, and the designed MR devices can thus be applied to other species with visual perception in both academic and industrial scenarios. For example, future research and application on enhancing cows’ perception of living on open grassland with MR devices to increase milk production and quality could be inspired by this work. Enhancing the perception of disabled people with MR devices is another possible application. However, for organisms that lack visual perception, such as some insects (e.g. ants) and plants, other perceptions need to be considered for enhancement to build a cross-species interaction biohybrid system.

These experiments have investigated the impacts of different enhanced perceptions (situation- and communication-oriented perceptions) on increasing people’s attention and promoting social understanding of natural organisms. However, the experiments are not perfect, and there is still some ongoing work to complete and analyze with this benchmark. Further studies on how perception-enhanced human–nature interaction can influence the cognition and behavior of natural organisms can be investigated (e.g. with fish). From the perspective of natural fish, the interaction between fish and artificial systems in this experiment is only performed through virtual ‘fish’ images on the back screen of the tank, which only provides the living fish with additional visual stimulus. The visual stimulus may not be captured by the fish during experiments due to their poor eyesight. In the future, biomimetic robot fish could be used to replace virtual ‘fish’ images, providing natural fish with more aspects of enhanced perception when interacting with humans. From the aspect of user experience, all the computation and objective detection algorithms can be integrated into MR devices (HoloLens) without additional screens to enhance the experiment’s reality and users can be better immersed in the biohybrid system.

Data availability statement

The data that support the findings of this study are openly available at the following URL/DOI: http://58.246.144.58:19700/d/710c4e9d398e45558a90/.

Acknowledgment

This work is supported by Shanghai Technology Development and Entrepreneurship Platform for Neuromorphic and AI SoC and in part by the Shanghai Research and Innovation Functional Program under Grant No. 17DZ2260900.

Ethics statement

This research follows the guidelines from the Technical Committee for Laboratory Animal Sciences of the Standardisation Administration of China (SAC/TC281). All treatments of experimental animals complied with the laws of the country (China) in which they were performed. All experiments are behavioral observations, and no permits are required in the country where the experiments are conducted. Before the test phase, having five animals per tank was not considered stressful, since this is a group-living fish. No fishes were injured or killed during the experiments.

Appendix A. Subjective surveys following scenes

A.1. Background
(a) Gender
(b) Age
(c) Have you ever raised pets?
1. Yes, which kind of pet?
2. If no, do you want to raise pets?
(i) If yes, which kind of pet?
(ii) No.
(d) Have you ever experienced VR/AR?
1. Yes
2. No

A.2. Scene null
(a) In this scene, which one of the following attracts your attention most?
1. Goldfish
2. Rocks
3. Others
(b) In this scene, I pay attention to these goldfish.
   1. Strongly disagree
   2. Disagree
   3. Neutral
   4. Agree
   5. Strongly agree

(e) In this scene, these goldfish react to my behavior.
   1. Strongly disagree
   2. Disagree
   3. Neutral
   4. Agree
   5. Strongly agree

(c) In this scene, I want to interact with these goldfish.
   1. Strongly disagree
   2. Disagree
   3. Neutral
   4. Agree
   5. Strongly agree

(f) In this scene, I can understand the mood of these goldfish.
   1. Strongly disagree
   2. Disagree
   3. Neutral
   4. Agree
   5. Strongly agree

(g) I want to have the ability to know the action trail of animals in the future.
   1. Strongly disagree
   2. Disagree
   3. Neutral
   4. Agree
   5. Strongly agree

A.3. Scene trail

(a) In this scene, which one of the following attracts your attention most?
   1. Goldfish
   2. Trails
   3. Rocks
   4. Others

(b) In this scene, I pay attention to these goldfish.
   1. Strongly disagree
   2. Disagree
   3. Neutral
   4. Agree
   5. Strongly agree

(c) In this scene, I want to interact with these goldfish.
   1. Strongly disagree
   2. Disagree
   3. Neutral
   4. Agree
   5. Strongly agree

(d) In this scene, these goldfish understand my behavior.
   1. Strongly disagree
   2. Disagree
   3. Neutral
   4. Agree
   5. Strongly agree

A.4. Scene mood

(a) In this scene, which one of the following attracts your attention most?
   1. Goldfish
   2. Emojis
   3. Rocks
   4. Others

(b) In this scene, I pay attention to these goldfish.
   1. Strongly disagree
   2. Disagree
   3. Neutral
   4. Agree
   5. Strongly agree

(c) In this scene, I want to interact with these goldfish.
   1. Strongly disagree
   2. Disagree
   3. Neutral
   4. Agree
   5. Strongly agree

(d) In this scene, these goldfish understand my behavior.
   1. Strongly disagree
   2. Disagree
   3. Neutral
   4. Agree
   5. Strongly agree

(e) In this scene, these goldfish react to my behavior.
   1. Strongly disagree
   2. Disagree
   3. Neutral
   4. Agree
   5. Strongly agree
In this scene, I can understand the mood of these goldfish.
1. Strongly disagree
2. Disagree
3. Neutral
4. Agree
5. Strongly agree

I want to have the ability to know the mood of animals in the future.
1. Strongly disagree
2. Disagree
3. Neutral
4. Agree
5. Strongly agree

A.5. Scene M&T
(a) In this scene, which one of the following attracts your attention most?
1. Goldfish
2. Trails
3. Emojis
4. Stone
5. Others
(b) In this scene, I pay attention to these goldfish.
1. Strongly disagree
2. Disagree
3. Neutral
4. Agree
5. Strongly agree
(c) In this scene, I want to interact with these goldfish.
1. Strongly disagree
2. Disagree
3. Neutral
4. Agree
5. Strongly agree
(d) In this scene, these goldfish understand my behavior.
1. Strongly disagree
2. Disagree
3. Neutral
4. Agree
5. Strongly agree
(e) In this scene, these goldfish react to my behavior.
1. Strongly disagree
2. Disagree
3. Neutral
4. Agree
5. Strongly agree
(f) In this scene, I can understand the mood of these goldfish.
1. Strongly disagree
2. Disagree
3. Neutral
4. Agree
5. Strongly agree

Appendix B. Hypothesis-testing

Table B1. Hypothesis-testing of accumulative attention.

| Test | Scene | p-value |
|------|-------|---------|
| Friedman | NULL, RT1, RT2, RT3 | 0.0061<sup>b</sup> |
| Wilcoxon | NULL vs RT1 | 0.0077<sup>b</sup> |
| Wilcoxon | NULL vs RT2 | 0.0008<sup>c</sup> |
| Wilcoxon | NULL vs RT3 | 0.0456<sup>a</sup> |
| Wilcoxon | RT1 vs RT2 | 1 |
| Wilcoxon | RT1 vs RT3 | 0.0524 |
| Wilcoxon | RT2 vs RT3 | 0.0351<sup>a</sup> |
| (b) NULL, MT1, MT2, MT3 | | |
| Friedman | NULL, MT1, MT2, MT3 | 0.0001<sup>c</sup> |
| Wilcoxon | NULL vs MT1 | 0.0018<sup>b</sup> |
| Wilcoxon | NULL vs MT2 | 0 <sup>c</sup> |
| Wilcoxon | NULL vs MT3 | 0.0014<sup>b</sup> |
| Wilcoxon | MT1 vs MT2 | 0.3272 |
| Wilcoxon | MT1 vs MT3 | 0.9844 |
| Wilcoxon | MT2 vs MT3 | 0.2027 |
| (c) RT3, MT3, RT3&MT3 | | |
| Friedman | RT3, MT3, RT3&MT3 | 0.1310 |
| Wilcoxon | RT3 vs MT3 | 0.0919 |
| Wilcoxon | RT3 vs RT3&MT3 | 0.0882 |
| Wilcoxon | MT3 vs RT3&MT3 | 0.6664 |
| (d) RT1, MT3, RT1&MT3 | | |
| Friedman | RT1, MT3, RT1&MT3 | 0.8669 |
| Wilcoxon | RT1 vs MT3 | 0.6827 |
| Wilcoxon | RT1 vs RT1&MT3 | 0.9729 |
| Wilcoxon | MT3 vs RT1&MT3 | 0.8382 |
| (e) RT1, MT1, RT1&MT1 | | |
| Friedman | RT1, MT1, RT1&MT1 | 0.8007 |
| Wilcoxon | RT1 vs MT1 | 0.8650 |
| Wilcoxon | RT1 vs RT1&MT1 | 0.4171 |
| Wilcoxon | MT1 vs RT1&MT1 | 0.2837 |
| (f) RT3, MT1, RT3&MT1 | | |
| Friedman | RT3, MT1, RT3&MT1 | 0.4378 |
| Wilcoxon | RT3 vs MT1 | 0.4452 |
| Wilcoxon | RT3 vs RT3&MT1 | 0.4634 |
| Wilcoxon | MT1 vs RT3&MT1 | 0.9406 |

<sup>a</sup> Significant at 5% level.
<sup>b</sup> Significant at 1% level.
<sup>c</sup> Significant at 0.1% level.
### Table B2. Hypothesis-testing of number of interactions.

| Test          | Scene               | p-value |
|---------------|---------------------|---------|
| Friedman      | Null, Trail, Mood, M&T | 0.7343  |
| Wilcoxon      | Null vs Trail       | 0.8201  |
| Wilcoxon      | Null vs Mood        | 0.2131  |
| Wilcoxon      | Null vs M&T         | 0.7535  |
| Wilcoxon      | Trail vs Mood       | 0.2490  |
| Wilcoxon      | Trail vs M&T        | 0.5809  |
| Wilcoxon      | Mood vs M&T         | 0.2013  |

### Table B3. Hypothesis-testing of subjective surveys.

(a) Results of A.2.2, A.3.2, A.4.2, A.5.2 in appendix A

| Test          | Questions                      | p-value |
|---------------|--------------------------------|---------|
| Friedman      | A.2.2, A.3.2, A.4.2, A.5.2       | 0.3568  |
| Wilcoxon      | A.2.2 vs A.3.2                  | 0.2708  |
| Wilcoxon      | A.2.2 vs A.4.2                  | 0.7963  |
| Wilcoxon      | A.2.2 vs A.5.2                  | 0.4913  |
| Wilcoxon      | A.3.2 vs A.4.2                  | 0.2359  |
| Wilcoxon      | A.3.2 vs A.5.2                  | 0.5367  |
| Wilcoxon      | A.4.2 vs A.5.2                  | 0.3173  |

(b) Results of A.2.3, A.3.3, A.4.3, A.5.3 in appendix A

| Test          | Questions                      | p-value |
|---------------|--------------------------------|---------|
| Friedman      | A.2.3, A.3.3, A.4.3, A.5.3       | 0.0329a |
| Wilcoxon      | A.2.3 vs A.3.3                  | 0.0048b |
| Wilcoxon      | A.2.3 vs A.4.3                  | 0.1853  |
| Wilcoxon      | A.2.3 vs A.5.3                  | 0.6647  |
| Wilcoxon      | A.3.3 vs A.4.3                  | 0.2866  |
| Wilcoxon      | A.3.3 vs A.5.3                  | 0.0807  |
| Wilcoxon      | A.4.3 vs A.5.3                  | 0.3841  |

(c) Results of A.2.4, A.3.4, A.4.4, A.5.4 in appendix A

| Test          | Questions                      | p-value |
|---------------|--------------------------------|---------|
| Friedman      | A.2.4, A.3.4, A.4.4, A.5.4       | 0.6361  |
| Wilcoxon      | A.2.4 vs A.3.4                  | 0.8610  |
| Wilcoxon      | A.2.4 vs A.4.4                  | 0.6025  |
| Wilcoxon      | A.2.4 vs A.5.4                  | 0.2535  |
| Wilcoxon      | A.3.4 vs A.4.4                  | 0.8332  |
| Wilcoxon      | A.3.4 vs A.5.4                  | 0.3297  |
| Wilcoxon      | A.4.4 vs A.5.4                  | 0.5046  |

(f) Results of A.3.7, A.4.7 in appendix A

| Test          | Questions                      | p-value |
|---------------|--------------------------------|---------|
| Wilcoxon      | A.3.7 vs A.4.7                  | 0.0235a |

### Table B3. (Continued.)

(d) Results of A.2.5, A.3.5, A.4.5, A.5.5 in appendix A

| Test          | Questions                      | p-value |
|---------------|--------------------------------|---------|
| Friedman      | A.2.5, A.3.5, A.4.5, A.5.5       | 0.1589  |
| Wilcoxon      | A.2.5 vs A.3.5                  | 0.0605  |
| Wilcoxon      | A.2.5 vs A.4.5                  | 0.7777  |
| Wilcoxon      | A.2.5 vs A.5.5                  | 0.9362  |
| Wilcoxon      | A.3.5 vs A.4.5                  | 0.0843  |
| Wilcoxon      | A.3.5 vs A.5.5                  | 0.0249a |
| Wilcoxon      | A.4.5 vs A.5.5                  | 0.6643  |

(e) Results of A.2.6, A.3.6, A.4.6, A.5.6 in appendix A

| Test          | Questions                      | p-value |
|---------------|--------------------------------|---------|
| Friedman      | A.2.6, A.3.6, A.4.6, A.5.6       | 0.0002c |
| Wilcoxon      | A.2.6 vs A.3.6                  | 0.4898  |
| Wilcoxon      | A.2.6 vs A.4.6                  | 0.0378a |
| Wilcoxon      | A.2.6 vs A.5.6                  | 0.0124a |
| Wilcoxon      | A.3.6 vs A.4.6                  | 0.0046b |
| Wilcoxon      | A.3.6 vs A.5.6                  | 0.0031  |
| Wilcoxon      | A.4.6 vs A.5.6                  | 0.7270  |

### References

Adolphs R 2006 How do we know the minds of others? Domain-specificity, simulation and enactive social cognition *Brain Res.* 1079 25–35

Azhar S, Kim J and Salman A 2018 Implementing virtual reality and mixed reality technologies in construction education: students’ perceptions and lessons learned 11th Annual Int. Conf. of Education (https://doi.org/10.21125/iceri.2018.0183)

Bates L A, Sayialel K N, Nijiraini N W, Moss C J, Poole J H and Byrne R W 2007 Elephants classify human ethnic groups by odor and garment color *Curr. Biol.* 17 1938–42

Bijlanka K and Loeschcke V 2013 Environmental Stress, Adaptation and Evolution (Basel: Birkhäuser)

Bischoff R and Graefe V 1999 Integrating vision, touch and natural language in the control of a situation-oriented behavior-based humanoid robot *Proc. 1999 IEEE Int. Conf. on Systems, Man and Cybernetics* vol 2 (IEEE) pp 999–1004

Bodenhansen G V and Hugenberg K 2009 Attention, perception and social cognition Social Cognition: The Basis of Human Interaction (Milton Park: Taylor and Francis) pp 1–22

Bruner J and Postman L 1949 Perception, cognition and behavior *J. Personality* 18 14–31
Byrne C, Freil L, Starner T and Jackson M 2017 A method to evaluate haptic interfaces for working dogs Int. J. Hum.-Comput. Stud. 98 196–207

Cahill J F Jr Bao T, Maloney M and Kolomosky C 2013 Mechanical leaf damage causes localized, but not systemic, changes in leaf movement behavior of the sensitive plant, Mimosa pudica (Fabaceae) J. Botany 91 43–47

Chen X, Duan M, Xie Q, Lai Y, Dong L, Cao W, Yao D and Luo C 2015 Functional disconnection between the visual cortex and the sensorimotor cortex suggests a potential mechanism for self-disorder in schizophrenia Schizophrenia Res. 166 151–7

Cheneler D, Buselli E, Camboni D, Anthony C, Grover L, Adams M J and Oddo C M 2014 A bio-hybrid tactile sensor incorporating living artificial skin and an impedance sensing array Sensors 14 23781–802

Chu Y, Li X, Yang X, Ai D, Huang Y, Song H, Jiang Y, Wang Y, Chen X and Yang J 2018 Perception enhancement using importance-driven hybrid rendering for augmented reality based endoscopic surgical navigation Biomed. Opt. Express 9 5205–26

Clark F E, Gray S I, Bennett P, Mason L J and Burgess K V 2019 High-tech and tactile: cognitive enrichment for zoo-housed gorillas Frontiers Psychol. 10 1574

Clifton G T, Bolton W and Gravish N 2020 Vision does not impact walking performance in argentine ants J. Exp. Biol. 223 jeb228460

Condino S, Turini G, Parchi P D, Vigliarolo R M, Pialonti N, Gesi M, Ferrari M and Ferrari V 2018 How to build a patient-specific hybrid simulator for orthopaedic open surgery: benefits and limits of mixed-reality using the Microsoft HoloLens J. Healthc. Eng. 2018 1–12

Corbetta M and Shulman G L 2002 Control of goal-directed and stimulus-driven attention in the brain Nat. Rev. Neurosci. 3 201–15

De Tinguay X, Pacchierotti C, Marchal M and Lécuyer A 2018 Enhancing the stiffness perception of tangible objects in mixed reality using wearable haptics 2018 IEEE Conf. on Virtual Reality and 3D User Interfaces (VR) (IEEE) pp 81–90

Esch H E, Zhang S, Sririvasan M V and Tautz J 2001 Honeybee dances communicate distances measured by optoelectronics Nature 411 581–3

Evans L T 1996 Crop Evolution, Adaptation and Yield (Cambridge: Cambridge University Press)

Goodman D A and Weinberger N M 1973 Amphibia as vertebrate model systems Habituation: Behav. Stud. 1 85

Griparić K, Haus T, Milčić D, Polić M and Bogdan S 2017 A robotic system for detecting social integration in honeybees PLoS One 12 e018197

Guo T, Zhuang L, Qin Z, Zhang R, Hu N and Wang P 2016 Multi-odor discrimination by a novel bio-hybrid sensing preserving rat’s intact smell perception in vivo Sens. Actuators B 225 334–41

Hallow J, Mondada F, Kernbach S and Schmickl T 2013 Towards bio-hybrid systems made of social animals and robots Conf. on Biomimetic and Biohybrid Systems (Springer) pp 384–6

Horne R B R 2012 Treatment perceptions and self-regulation The Self-Regulation of Health and Illness Behaviour (Milton Park: Routledge) pp 152–68

Ilgün A et al 2021 Bio-hybrid systems for ecosystem level effects ALIFE 2021: The 2021 Conf. on Artificial Life (MIT Press) (https://doi.org/10.1162/isal_a_00396)

Isokawa N, Nakazawa J and Okoshi T 2019 AffectiveNemo: promoting owner–fish interaction by displaying observed fish emotions with speech bubble Proc. of the 6th Int. Conf. on Animal-Computer Interaction pp 1–5

Jang J, Jun Y S, Seo H, Kim M and Park J K 2020 Motion detection using tactile sensors based on pressure-sensitive transistor arrays Sensors 20 3624

Kessel D, García-Rubio M J, González E K, Tapia M, López-Martín S, Román F J, Capilla A, Martínez K, Colom R and Carretié L 2016 Working memory of emotional stimuli: electrophysiological characterization Biol. Psychol. 119 190–9

Kihlstrom J F 1987 The cognitive unconscious Science 237 1445–52

Kim L, He L, Maaswinkel H, Zhu L, Sirotkin H and Weng W 2014 Anxiety, hyperactivity and stereotypy in a zebrafish model of fragile X syndrome and autism spectrum disorder Prog. Neuro-Psychopharmacology Biol. Psychiatry 53 46–49

Krsuljan E 2007 Fundamental components of attention Annu. Rev. Neurosci. 30 57–78

Ko D, Kwon D, Kim E and Lee W 2018 BubbleTalk: enriching experience with fish by supporting human behavior Proc. of the 2018 Designing Interactive Systems Conf. pp 919–30

Laming P and Savage G 1980 Physiological changes observed in the goldfish (Carassius auratus) during behavioral arousal and fright Behav. Neural Biol. 29 235–57

Landgraf T, Oertel M, Kirbach A, Menzel R and Rojas R 2012 Imitation of the honeybee dance communication system by means of a biomimetic robot Conf. on Biomimetic and Biohybrid Systems (Springer) pp 132–43

Landgraf T, Rojas R, Nguyen H, Kriegel F and Stettin K 2011 Analysis of the waggle dance motion of honeybees for the design of a biomimetic honeybee robot PLoS One 6 e21354

Lazic D and Schmickl T 2021 Can robots inform a honeybee colony’s foraging decision-making? ALIFE 2021: The 2021 Conf. on Artificial Life (MIT Press) (https://doi.org/10.1162/isal_a_00397)

Liu Q, Hu N, Zhang F, Zhang D, Hsia K J and Wang P 2012 Offactory epithelium biosensor: odor discrimination of receptor neurons from a bio-hybrid sensing system Biomed. Microdevices 14 1055–63

Logunov D and Konnov M 1983 Model investigation of neuronal mechanisms of discrimination between two tactile stimuli in snails Neurophysiology 15 445–50

Low Z W K, Li Z, Owch C, Chee P L, Ye E, Kai D, Yang D P and Loh X J 2019 Using artificial skin devices as skin replacements: insights into superficial treatment Small 15 1805453

Lucarotti C, Oddo C M, Vitello N and Carrozza M C 2013 Synthetic and bio-artificial tactile sensing: a review Sensors 13 1435–66

Mackie D M, Devos T and Smith E R 2000 Intergroup emotions: explaining offensive action tendencies in an intergroup context J. Personality Soc. Psychol. 79 602

Macrae C N and Bodenhausen G V 2001 Social cognition: categorical person perception Br. J. Psychol. 92 239–55

Michelsen A, Andersen B B, Storm J, Kirchner W H and Lindauer M 1992 How honeybees perceive communication dances, studied by means of a mechanical model Behav. Ecol. Sociobiol. 30 143–50

Nickell W T, Kleene N K and Kleene S J 2007 Mechanisms of neuronal chloride accumulation in intact mouse olfactory epithelium J. Physiol. 583 1005–20

Pepino M Y, Bradley D, Eagon C J, Sullivan S, Abumrad N A and Klein S 2014 Changes in taste perception and eating behavior after bariatric surgery-induced weight loss in overweight women Obesity 22 E13–E20

Popper A N and Carlson T J 1998 Application of sound and other stimuli to control fish behavior Trans. Am. Fisheries Soc. 127 673–707

Ratnieks F L 2007 Biomimicry: further insights from ant colonies? Workshop on Bio-Inspired Design of Networks (Berlin: Springer) pp 58–66

Rensink R A 2013 Perception and attention The Oxford Handbook of Cognitive Psychology (Oxford: Oxford University Press)

Richardson J L, Urban M C, Bolnick D I and Skelly D K 2014 Microgeographic adaptation and the spatial scale of evolution Trends Ecol. Evol. 29 165–76

Roberts K L and Allen H A 2016 Perception and cognition in the ageing brain: a brief review of the short-and-long-term links between perceptual and cognitive decline Front. Aging Neurosci. 8 39
Romano D, Donati E, Benelli G and Stefanini C 2019 A review on animal–robot interaction: from bio-hybrid organisms to mixed societies *Biol. Cybern.* **113** 201–25
Romano D and Stefanini C 2022a Any colour you like: fish interacting with bioinspired robots unravel mechanisms promoting mixed phenotype aggregations *Bioinspir. Biomim.* **17** 045004
Romano D and Stefanini C 2022b Robot–fish interaction helps to trigger social buffering in neon tetras: the potential role of social robotics in treating anxiety *Int. J. Soc. Robot.* **14** 963–72
Seidl T and Wehner R 2006 Visual and tactile learning of ground structures in desert ants *J. Exp. Biol.* **209** 3336–44
Smith E R and Mackie D M 2016 Group-level emotions *Curr. Opin. Psychol.* **11** 15–19
Smith J B and Alloway K D 2013 Rat whisker motor cortex is subdivided into sensory-input and motor-output areas *Front. Neural Circuits* **7** 4
Standing L, Conezio J and Haber R N 1970 Perception and memory for pictures: single-trial learning of 2500 visual stimuli *Psychon. Sci.* **19** 73–74
Stephenson L J, Edwards S G and Bayliss A P 2021 From gaze perception to social cognition: the shared-attention system *Perspect. Psychol. Sci.* **16** 553–76
Tyan Y-S, Wu M-C, Chin C-L, Kuo Y-L, Lee M-S and Chang H-Y 2014 Ischemic stroke detection system with a computer-aided diagnostic ability using an unsupervised feature perception enhancement method *Int. J. Biomed. Imaging* **2014** 1–12
Wolfe B, Sawyer B D and Rosenholtz R 2022 Toward a theory of visual information acquisition in driving *Hum. Factors* **64** 694–713
Zhang B, Zhuang L, Qin Z, Wei X, Yuan Q, Qin C and Wang P 2018 A wearable system for olfactory electrophysiological recording and animal motion control *J. Neurosci. Methods* **307** 221–9