Article

Allocentric Emotional Affordances in HRI: The Multimodal Binding

Jordi Vallverdú[1,*], Gabriele Trovato[2], and Lorenzo Jamone[3]

[1] Universitat Autònoma de Barcelona, Catalonia; Jordi.vallverdu@uab.cat
[2] Waseda University, Japan; gabriele@takanishi.mech.waseda.ac.jp
[3] Queen Mary University of London, UK; l.jamone@qmul.ac.uk

* Correspondence: jordi.vallverdu@uab.cat ; Tel.: +34-581-16-18

Abstract: Affordances are an important concept in cognition, which can be applied to robots in order to perform a successful human-robot interaction (HRI). In this paper we explore and discuss the idea of emotional affordances and propose a viable model for implementation into HRI. We consider "2-ways" affordances: perceived object triggering an emotion, and perceived human emotion expression triggering an action. In order to make the implementation generic, the proposed model includes a library that can be customised depending on the specific robot and application's scenario. We present the AAA (Affordance-Appraisal-Arousal) model, which incorporates Plutchik's Wheel of Emotions, and show some examples of simulation and possible scenarios.

Keywords: affordance, empathy, HRI, emotion, multimodal, allocentric, libraries.

1. Introduction: An Affordable World

In 1966, the psychologist James Gibson introduced a new concept [1], "affordances" , to define an action-centric perspective on the visual perception of the external environment. His best-known definition is taken from his seminal 1979 book, *The Ecological Approach to Visual Perception* [2]: "The affordances of the environment are what it offers the animal, what it provides or furnishes, either for good or ill. The verb to afford is found in the dictionary, the noun affordance is not. I have made it up. I mean by it something that refers to both the environment and the animal in a way that no existing term does. It implies the complementarity of the animal and the environment". As it became a fundamental concept during the next decades a vast literature about it meaning has being produced [3]–[6]. It is beyond the scope and aims of this paper to discuss in detail the very notion of affordance; however, we do explain in which context we apply it, and what is the special value of our model. Specifically, here we investigate how the concept of affordances, that was originally formalised in the context of visual perception of action-related properties of objects and spaces, can relate to *emotions*, and how could it be used in Human-Robot Interaction (henceforth, HRI) scenarios. Indeed, the use of affordances in HRI, can facilitate social interactions between humans and robots or, more generally, between humans and machines. Moreover, emotions are a fundamental cognitive and social aspect of social life [7], [8], and their value has also been considered of the outmost importance for the design of artificial cognitive architectures [9].

Through previous researches on emotional modelling [10]–[12] we have contributed to the design of a multimodal model for the use of emotional affordances in HRI [13], [14]. Most of HRI literature that includes emotional affordances is biased by two important aspects: first, it maintains an over-simplistic emotional mapping based on 6 Ekman's emotions[15], [16], and secondly it defends an egocentric spatial imagery, which involves a self-to-object representational system. Most of current studies on robotics and emotions are based on very basic aspects of emotional clues, mainly, visual ones (and based on a one-to-one direct connection with Ekman's thesis). Instead, here

we want to consider the temporal nature of emotional modulations as well as the multidimensional elicitators of emotional values..

We have previously discussed some of these issues [17][18], and with this paper we extend those discussions and provide a new way to consider emotional affordances in HRI. In addition, we propose a viable model for implementation into HRI.

We implement the case of 2-way affordances:

- Perceived object triggering an emotion
- Perceived human emotion expression triggering an action

In order to make the implementation generic and adaptable, the proposed model encapsulates a library that can be customised through robot-specific and scenario-specific tuneable parameters, that will have different initializations depending on the situation and will also be adapted online based on the feedback that the robot receives during its interaction with the environment and with the human users.

The rest of the manuscript is structured as follows: in Section 2, we introduce the background and the concepts related to affordances; in Section 3 we describe the model and show some examples of simulation in Section 4. In Section 5 we discuss possible scenario, and in Section 6 we summarise and conclude the paper.

2. Background

In the following subsections we motivate our attempt to integrate the concepts of emotions and affordances for a successful HRI, and we outline the fundamental aspects that characterise our model.

2.1. The grounded nature of emotional syntax

Not only emotions lie at the basis of any informational-cognitive event, but they become also a mechanism to manage internal data as well as to share information with other agents. In that sense, emotions are embodied [19] and are the result of a natural evolutionary process that has increased survival ratios among those entities which use them, because of the greater adaptive skills they make possible [20]–[22]. This embodied nature of emotions affects also the way by which we perceive and construct our emotional experiences [23].

Notably, Darwin [24] was the first one to think about the expression of emotions in humans (more precisely, only about "man", writing gendered biased) and animals from an evolutionary and naturalistic perspective. The idea is simple: the arousal of emotions provides information to other beings about our internal states and help to synchronise possible actions together. But the feeling about internal processes it helps us. Indeed, people locate and identify health states thanks to emotional awareness [25]. People with autism or Alexithymia, for example, are bad at identifying not only other's emotions but also their own ones, and as a consequence they are not able to identify their own health problems [26], [27]. Thus, introducing a functional perspective, emotions serve as identifiers and labellers of informational states individual as well as in social contexts. For that reason emotions are clues for the understanding of the state and intentions of an agent. Until now, most of roboticists and engineers have followed what we will call the ''Skinnerian approach to emotions''. Due to the complexities about emotional nature understanding and the myriad of divergent emotional models (and, obviously, the impossibility at this historical point to have true emotional machines, [9]), most experts have reduced their interests to observable external aspects of emotions. Despite of the utility of such over-simplistic approach for HRI or HCI, the fact is that this Skinnerian approach turns blind systems to the grounded causes that connect emotional states and manage at the same time their temporal syntax. The emotional dysfunction of our intelligent systems is deep, because it is not only necessary to capture basic external emotional signals, but also to

connect them to create a map of the grounded nature that produces them. Our empathy is not only a moral value but also a sensorimotor mechanism to copy actions and learn by imitation [28], [29].

2.2. Affordances and emotions

Although this paper does not aim to revisit the long debates on the meaning of 'affordance' (see [30], and [31] for a recent survey on affordance related studies in different disciplines), let us consider a general definition: an affordance is a relation between an informational event (objects, environments, even virtual or imaginary things) and an organism which, through the related stimuli and the heuristically processing of them, affords the opportunity for that organism to perform an action. Easily: it's the availability of possible uses of an object by an entity, like a cord affords pulling. Following original Gibson uses [3], affordances are resources that any animal use for some purpose. Thus, it is the animal who perceives something as an affordance, something 'useful for' or 'interesting to', and the affordance has an ecological nature. Although the performer of the experiences creates the use, it is also true that the correlated nature of the object/environment together with the expectations, skills and intentionality of the user (both actual and perceived properties) are what make the emergence of affordances possible.

Notably, emotions set the context for action selection, hierarchically ordered following biological priorities [32], [33]. In this sense, as stated in [34], emotions lie at every core goal of a living entity. Emotions are then not just behavioural external actions that need to be scanned and interpreted for secondary purposes of human actions analysis. This would justify an over-simplistic Skinnerian approach to emotional expression (while the semantic values that produce emotional expression would remain hidden). In order to avoid such bias into our models (both as a result of embodied and/or social neglections), we need to give emotions a more fundamental role: to be the basic indicators of the deep architecture and motivations that regulate human activities. Emotions ground and regulate any kind of informational interaction with the environment. Consequently, any approach to cognitive systems must consider them not because of their plain behaviourist outcome, but also for their role into the design of action responses.

Therefore, from a functional perspective, we can say that *emotions are affordances,* or that the process of feeling emotions is related to that of perceiving affordances, since both processes *modulate* the immediate sensing and *prepare* the agent for some possible future actions.

2.3. Emotional affordances of objects

Objects own emotional values, thanks to a transfer process made by human agents [35]. During this process humans even add and transfer anthropomorphic evaluations to objects (and to robots) [36]–[38]. For such reason, any serious emotional affordance library should not only include human bodily aspects related to emotional activities, but also a way to connect emotional values to the emotional actions in which humans tag those objects with emotional values. Therefore, our model includes (see Figues 1 and 2) the capture and combined evaluation of objects and emotional expressions, using a global multimodal approach, which is related to an allocentric analysis of human agency.

2.4. Physical affordances of emotions

This is the classic approah to emotional analysis: to consider faces and other related basic bodily expresions of emotions [24], [39], [40]. Despite of the possible debates about the universality or culturallly-related aspect of emotions elicitation, the basic aspect of these approaches is to understand how these models capture external emotional expressions and connect them by linear models to emotional states. In our model we will use them as well but as part of complex behavioural and dynamical contexts. A Skinnerian approach to emotions (one-to-one direct relationships between expressions and emotional descriptions) does not provide the best model for

a complete HRI process: the robot must be able to understand the natural temporary and emotionally transitions between moods and how the emotional states follow a coherent and even predictable dynamical evolution. There are fundamental biochemical mechanisms that explain these modulatory processes, like neuromodulatory ones [41]: in our robotic model we do not need a detailed representation of the physical mechanisms that make emotions possible in humans, but we do include the dynamical analysis of emotional variations to support natural and effective HRI.

2.5. Allocentric affordances

Although it is true that an affordance is processed by an agent inside an ecological niche, it has been a   common mistake in the HRI literature to focus only on a egocentric (subject-object) spatial analysis, mainly because of the laboratory experimental restrictions. Real human actions are performed in rich environments, which involve an object-to-object spatial relationships as well, and present therefore the need to encode information about the location of one object or its parts with respect to other objects objects (i.e. allocentric spatial analysis). This is clear when we analyse cultural psychologists who have checked the visualization patterns of different communities [42], [43]. It is this enriched social scenario in which emotional affordances are processed, and because of it we can understand how HRI can be manipulated in order to obtain agents responses (which are not purely automated) [44] .

2.6. Multimodal Binding

Consequently, our ontological approach includes (See **Figure 1)** an affordance phase in which objects and emotions are processed in a separated but at the same time connected way. Beyond the oversimplistic Ekmans model, we suggest a more precise Plutchik's wheel [45] that enables a more comprehensive emotional mapping in which emotions are also classified by intensity. At the same time, we consider the object classification process, which also includes a second problem: the discretization of actions in order to reconstruct the logic of the uses of objects in the work spaces. Thanks to an allocentric perspective, both input interface modules can be combined to obtain reliable information.

2.7. Dynamic processing and Libraries

A fundamental aspect of our model is the dynamical nature of libraries which must combine departing emotions with concluding related actions. A combination of parametric emotion recognition with their connection with the value of the object inside a taxonomy of recognised objects provides a diadic affordance appraisal mechanism: first, the output performance module is able to connect variable and dynamic emotional states inside a holistic framework in which the *answer is also emotionally-valued*; second, the motor tasks can be calculated according to possible aims.

2.8. A computational model for effective emotional binding in HRI

Therefore, summarising our main contributions, the allocentric, multimodal, and dynamic nature of our model makes a real emotional binding between robots and humans possible. Social robots that interact with humans (who at their turn also are socially bonded with other humans) need to understand not only the emotional data from human beings but how to connect them with the affordability of their available set of objects and the expected value of their related final actions. The arousal phase is also part of the expected outcome planning to-be-performed, because emotional processes imply not only the immediate physical response but also the evaluation of the elicitation mechanisms that explain and attribute meaning to agents' actions.

3. AAA model

We introduce the AAA (Affordance-Appraisal-Arousal) model, which generalises several mental processes happening in humans, and translates them to the case of robots. These three phases happen subsequently in the process that goes from the perception to the behavioural response. In Figure 1, the perceived sensorial feedback goes through a neural mechanism that generates affordance perception. Since the detailed description of the neural mechanisms that produce affordances perception in living organism's brain is not in the scope of this research, we leave it as an empty box, of which we know the inputs (some multimodal perception) and the possible outputs (an emotion display or a physical action, suggested by the affordance mechanism).

In Figure 1, the boundaries of the human or of the living organism are marked in the pink box. Out of it, we can observe the external events and the feedback to those events. The diagram is divided in three phases, namely Affordance, Appraisal and Arousal.
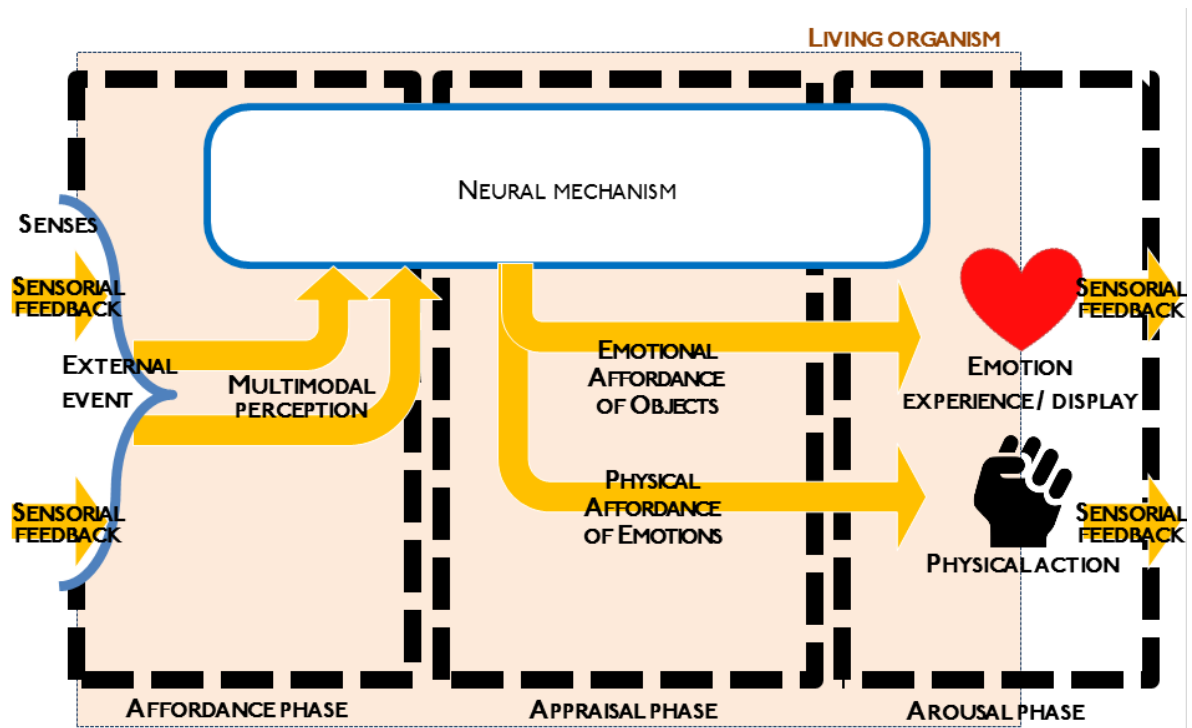

Figure 1. Affordance-Appraisal-Arousal model applied to humans or more generically to living organisms

For robots, we apply the same structure, except that the area marked in the light blue box in Figure 2 consists of the components of the library, which interfaces with the robot (in dark grey).
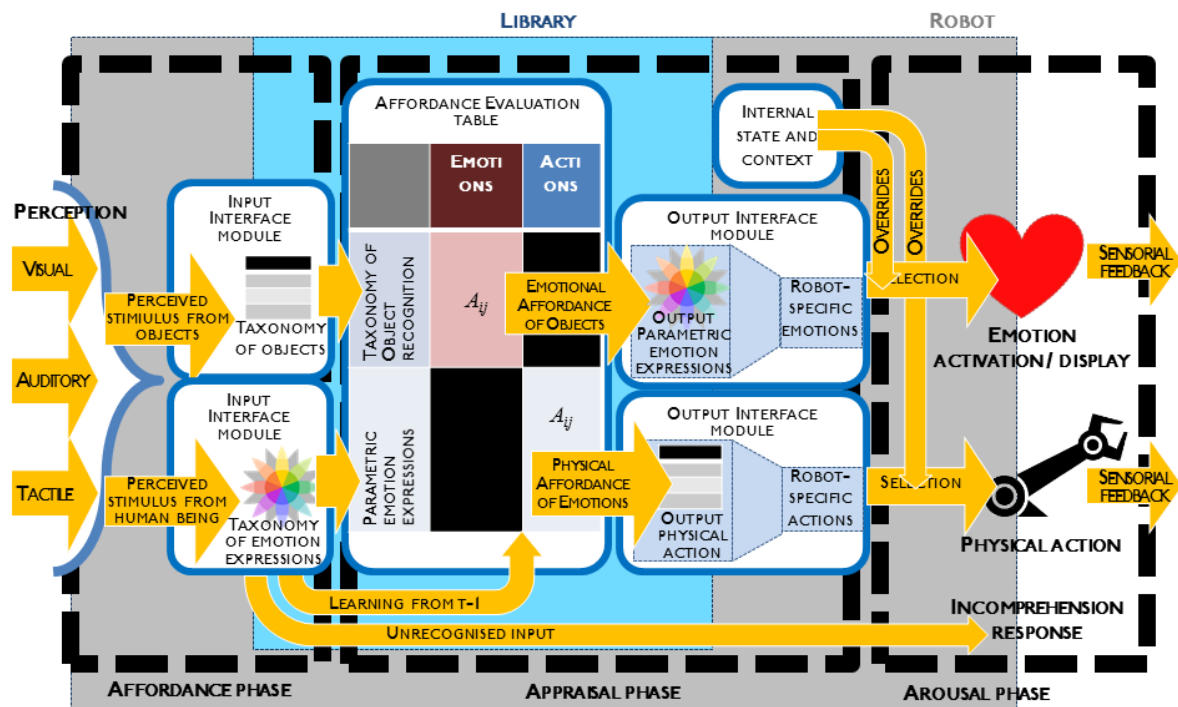
Figure 2. Affordance-Appraisal-Arousal model applied to robots

In the following subsections, we describe the content of the modules shown in Figure 2.

3.1. Perception

Robots typically sense the environment through different sensory channels, most commonly visual, auditory and tactile. The visual channel is the most commonly associated with affordances [46], [47]. In our model, we examine the generic case of a multimodal input, without entering deep into the analysis of the symbol grounding problem. We take for granted that a generic robot is able to perceive the environment and classify a certain number of stimuli.

We divide these stimuli in two broad categories: objects and humans. In the latter case, we focus specifically on human emotion expressions.

3.2. Input interfaces

The possible inputs that the robot can recognise must be interfaced with the taxonomies in the library.

This mapping is left here as a stub (in software development it is a piece of code used to stand in for some other programming functionality), containing some simple mapping of inputs associated with the library counterpart. The mapping can be customised. There are two interfaces: one for the objects, one for human expressions.

In the former case, the mapping is made by simple correspondance opearated by the function f on the raw input P_O. For example:
- "cup" -> "cup" (direct correspondance)
- "mug" -> "cup" (when the library taxonomy only has a similar entry)
- "wrench" -> null (when the library taxonomy does not have such entry)
- null -> "spoon" (when the robot does not recognise such object)

The class memberships to each label, in the form of an array, get copied when there is a correspondence, or set to zero in case of null correspondence. Subsequently, only the highest value is kept, while all the others are set to zero. The resulting data structure is also an array, which contains the class memberships to the labels of the objects available in the library. Within them, only one is set as different from zero. Of the resulting vector, only the element with the maximum value is kept, and the rest is set to 0 (equation 1), resulting in I_O (equation 2).

$$g(P_{O_i}) = \begin{cases} P_{O_i} & \text{for max } P_{O_i} \\ 0 & \text{otherwise} \end{cases} \tag{1}$$

$$I_O = g(f(P_O)) \tag{2}$$

In case of human expressions, the process is slightly more complicated, as it involves the use of an emotional model to parametrise the emotional content.

For this purpose, we use Plutchik's Wheel of Emotions [22], shown in Figure 3a. Plutchik's model is particularly convenient because it features emotions that are specular: anger is the opposite of fear; disgust is the opposite of trust, and so on. In a previous research [48], each of these emotions pair was referenced by an axis. By using 4 axis (one for each pair of leaves of Plutchik's Wheel, as in Figure 3b), the emotion can be parametrised. The four axis are called Mood, Stance, Temperament and Expectation, coherently with [48].

The interface in the case of human emotion expressions is therefore containing a mapping p between the raw input expressions P_E and 4 parameters, each one bounded between -1 and 1, resulting in I_E (equation 3). Example:

- "Smile" -> (0.5, 0, 0, 0)
- "Cry" -> (-1, 0, 0, 0)

The mapping does not convert all the inputs into emotions: it is executed only on the input with the maximum value among those which have a correspondence in the library ($\max f(P_E)$).

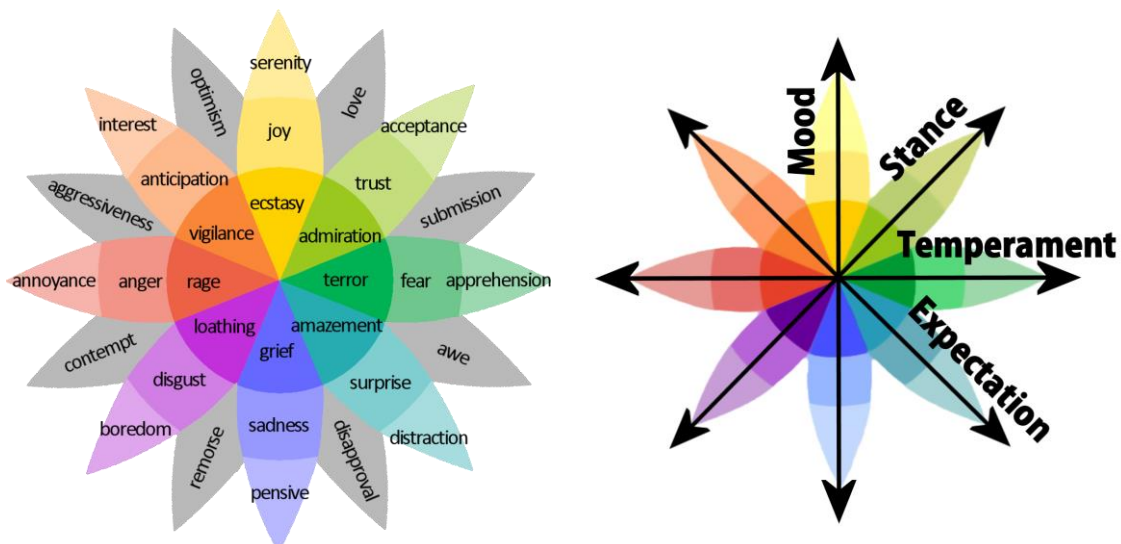$$I_E = p(\max f(P_E)) \tag{3}$$

Figure 3. Plutchik's Wheel of emotions (a) and its parametric implementation (b). Conceptually, in (b) the higher intensity of each emotion correspond to the tip of the leaves, with the zero located in the centre, rather than featuring the highest intensity in the centre, like in (a)

3.3. Affordance evaluation table

Table 1. Affordance Evaluation Table.

Title 1	Emotions	Actions
Taxonomy of object recognition	Case A: Emotional affordance of objects	Case B: Affordance of objects
Taxonomy of parametric emotion expressions	Case C: Empathy	Case D: Physical affordance of emotions

Case A: from an input consisting in the perception of an object, the emotional affordance selects a suggested emotion according to the formula:

$$R_E = I_O \cdot A_{OE} \qquad (4)$$

Where I_O is the row vector consisting in the input of n objects recognition class membership; A_{OE} is a matrix n x 4 consisting in the mapping; and R_E is the output, a row array of the 4 Plutchik's parameters.

This approach is similar to the sensing personality table made for the robot WE-4 [49]It contained a multimodal input of stimuli and their association to positive, neutral or negative activation, pleasentness and certainty.

Case B: perceiving the affordance of objects is a topic that has been studied in numerous previous researches [6]. Here we focus only on emotion affordances, as aspect that has been so far overlooked in the literature.

Case C: when an emotion expression triggers an emotion in the other party, it is a case of empathy [50], [51]. Empathy is the capacity to understand what another person is feeling, as is captured by popular wisdom with expressions like "put oneself in someone's shoes". It also encodes basic sensorimotor processes of learning and its study in Human-Robot Interactions has been widely studied [44], [52]. For this reason, we do not add this part to our affordance model.

Case D: from an input consisting in the perception of a human emotion expression, the emotional affordance selects a suggested action according to the formula:

$$R_A = I_E \cdot A_{EA} \qquad (5)$$

Where I_E is the row vector consisting in the input of the 4 components of Plutchik's parametrical model; A_{EA} is a matrix 4 x m consisting in the mapping; and R_A is the output, a row array of the m output action class memberships.

3.3.1. Learning

At any given time T, The tables A get updated at each input I_E. The values of A affected are the ones related to the latest (T-1) emotion expression or action executed by the robot. The valence v of human reaction is categorised according to the following formula:

$$v = \sum_l v_l \qquad (6)$$

Where l is the emotion in each leaf of Plutchik's model. By default, a positive valence is associated to Happiness, Trust. Negative valence is associated to Sadness, Disgust, Anger and Fear.

The resulting v acts as modifier, together with the learning rate α (bounded between 0 and 1), the element of A related to j, given that the robot performed the jth action or emotion expression, subsequently to the input i.

$$A_{i,j}{}^{T} = A_{i,j}{}^{T-1} \cdot (1 + \alpha \cdot v) \qquad (7)$$

The resulting afforace table evaluation will be boosted or reduced, according to human reaction.

3.3.1. Customisation

One possibility offered by this library is to customise the robot's affordance behaviour for different users or group of users. This will make the affordance mechanism flexible and applicable to a range of scenarios and users, each with different preferences (Figure 4).
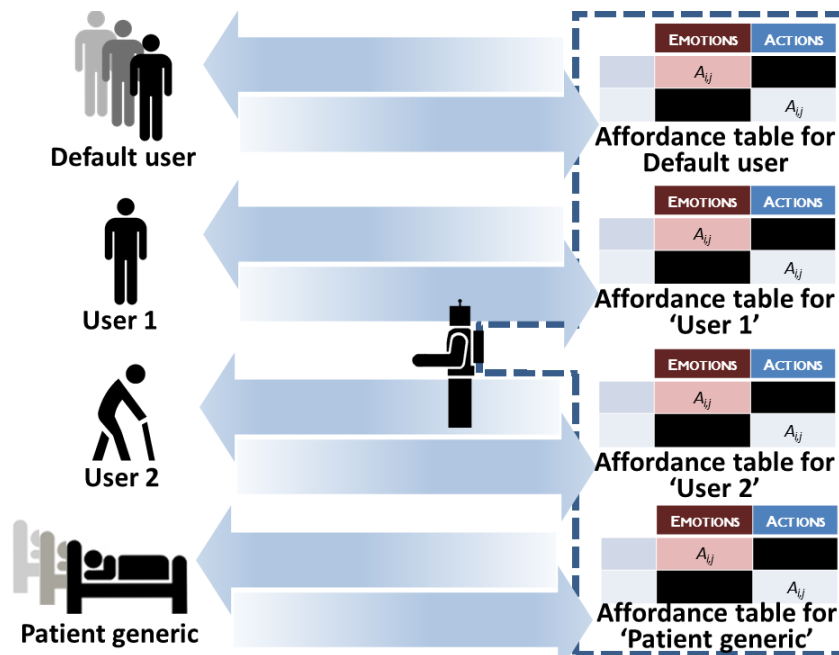

Figure 4. Concept of customised affordance evaluation tables

3.4. Output interfaces

The output interfaces work in a similar way to the input ones. Their role is to convert the output of the affordance evaluation table to the specific robot. There are two interfaces: one for the output emotions, one for the output actions.

The emotions as obtained by the previous step are a vector of 4 values, which can be positive or negative. They are converted into a vector of membership to an extensive list of all the emotions present in Plutchik's (function e in equation 8), and subsequently undergo a process of filtering (function f in equation 8), in order to match the robot's specific emotion capabilities. For example, if the robot is based on Ekman's standard 6 emotions, there will be a specific Plutchik-to-Ekman conversion process.

The suggested output of emotions S_E will be the result of these operations:

$$S_E = f\big(e(R_E)\big) \qquad (8)$$

On the other hand, the *m* output actions of S_A will feature positive or negative values, due to the previous matricial product involving the emotional parameters. For this reasons, the first step consists in setting to zero all the negative values as specified in the activation function *a* of equation 9:

$$a\left(R_{A_j}\right) = \begin{cases} 0 & \text{for } R_{A_j} < 0 \\ R_{A_j} & \text{for } R_{A_j} \geq 0 \end{cases} \tag{9}$$

Each membership R_{A_j} of an action *j* among the *m* in the library will be filtered in the same way as it happens with emotion, in order to match robot's action capabilities.

The suggested output of objects S_A will be the result of these operations:

$$S_A = f\big(a(R_A)\big) \tag{10}$$

3.5. Internal state and context

The internal state of the robot, and the degree of understanding the context, are specific of robot implementation. This part, necessary for reasoning, may contain its own emotional model and other cognitive constructs. Our affordance model is intended to work in parallel to robot reasoning and provide a faster feedback. The implementation of the robot can decide to let the result of the affordance reach the arousal phase, or to use that data for other calculations, or to completely override it.

3.6. Arousal

This part is the output of the model. It represents the execution of the emotion expression (as consequence of an affordance of an object) or the execution of an action (as consequence of an affordance of a human emotion expression). An additional action or emotion expression can be triggered in the case of incomprehension.

Generically speaking, the arousing emotion or action is determined by the maximum value of the vector produced by one of the output interfaces. These actual outputs can be overridden by the robot internal state, which is not part of the library. As the arousal is robot-specific, in this paper we do not focus on its implementation.

3.7. Case of incomprehension

In one of the following cases:
- When the raw input gets classified as "unknown" (either unknown object or unknown human expression)
- When each value of class membership of P_E or P_O is below a user-defined threshold
- When each value of the filtered input I_E or I_O is below a user-defined threshold (possibly as result of a missing match)

an affordance of incomprehension is triggered.

Comprehension is a very important aspect of communication: according to Poggi et al. [53], [54], the performative (namely, the actual action of interaction) of a communication act is defined by six parameters, one of which is the certainty of the communication content. In humans, the degree of certainty also influences facial cues [49], [54]. In robots, this is sometimes neglected, as affective communication only takes into consideration standard emotions: one of the few exception was KOBIAN-R [48], featuring a facial expression with currugated eyebrows especially for the case of incomprehension.

In our proposed architecture, when incomprehension happens, the whole process of affordance does not take place, and the suggested affordance is just "incomprehension": this might then

generate a related emotion expression or responsive action, depending on the specific robot implementation.

4. Sample operation

In this section we show some examples of how affordance perception and behaviour generation can be produced through the AAA model, by using the stub of the implementation described in the previous section, with simulated data.

4.1. Emotional affordance of objects

Let us suppose that a robot is surrounded by several objects. Through the combination of visual, auditory and tactile sensing, it can produce a vector P_O of raw input consisting in the class memberships.

$P_O = [0.4, 0.01, 0.01, 0.9, ...]$, which refer to the labels {'white ball', 'plate', 'pen', 'glass of water', 'box', ...}.

In the input interface, the filtering process $f(P_O)$ results in a new vector $I_O = [0.9, 0.4, 0, 0, 0, ...]$, referring to the labels available in the library {'glass of water', 'ball', 'cake', 'power socket', 'gun', ...}. In some cases (glass of water), there was direct correspondence; in other cases (ball) there was a correspondence of a similar entry; in all the cases with 0s (cake, power socket, gun) there was no correspondence.

The next step consists in processing the affordance evaluation table.

I_O	glass of water	ball	cake	power socket	gun	...
	0.9	0	0	0	0	

A_{OE}	mood	stance	temper ament	expect action
glass of water	0	-1	0.5	0
ball	0.5	0	0	0
cake	0.5	0	0	0.5
power socket	0.5	0.5	0	0
gun	0	0	1	0
...				

R_E	0	-0.9	0.45	0

Figure 5. Matrix multiplication resulting in a vector containing parametric emotion

The resulting vector $R_E = [0, -0.9, 0.45, 0]$ has 4 components. In its extended form $e(R_E)$, the vector indicates $[0, 0, 0, 0, 0.45, 0, ...]$ for the labels {'ecstasy', 'joy', 'serenity', 'admiration', 'trust', 'acceptance', 'terror', 'fear', 'apprehension',...}. The maximum value corresponds to 'loathing' (the most intense form of disgust, with 0.9).

Let us now suppose that our robot can only display happiness, sadness, anger and fear. The filtering process $f(e(R_E))$ will eliminate the values referring to disgust or loathing (unless specified by the user in a customised output interface), and just keep the value related to fear.

$S_E = [0, 0, 0, 0.45]$ is the final output of the library. The robot can decide to follow the suggested emotional affordance of fear and arouse an expressive behaviour, or to override it. Figure 6 summarises the whole process.
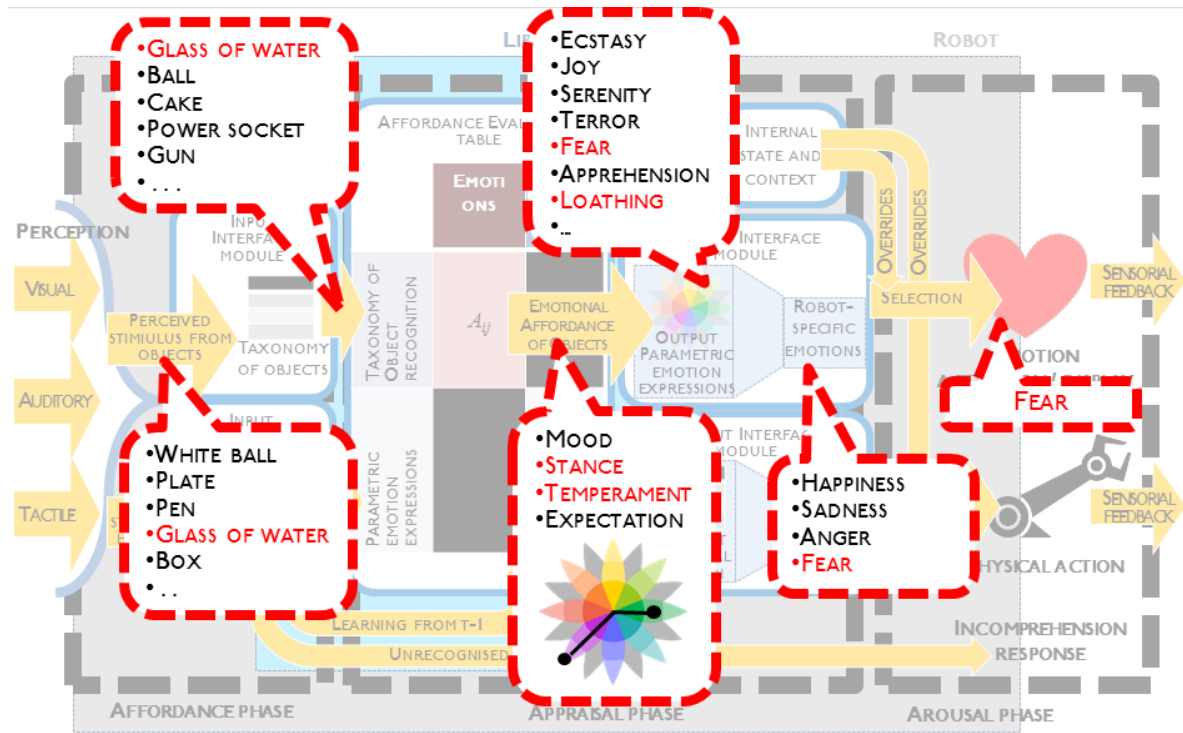
Figure 6. Summary of the object-to-emotion affordance process

4.2. Physical affordance of emotions

In this second example, the robot identifies a human face crying. The vector P_E of raw input contains [0.8, 0.1, 0.31, 0.6, 0.02 …], which refer to the labels {'cry', 'smile', 'head down', 'raise eyebrows', 'frown eyebrows', 'wrath', …};.

The filtering process results eliminates the entries that are not found in the library (i.e. 'wrath'), and in the resulting data, the maximum entry gets selected (cry, 0.8).

The mapping p evaluates the emotion of the current input parametrically. The resulting I_E for 'cry' is [−0.8, 0, 0, 0].

Figure 5 illustrates the product beween I_E and A_{EA}. The resulting vector R_A contains the affordance of actions. These actions contained in the library are mostly social actions, as they are supposed to be pertinent to respond to some human emotion expression.

Figure 7. Matrix multiplication resulting in a vector containing an action

However, the robot might not have all of these social actions implemented. The robot in our case can only perform the following actions: {'push', 'pull', 'take', 'greet', 'approach', 'hug'}. The filtering process results in $S_A = [0, \ 0, \ 0, \ 0, \ 0, 0.8]$. The whole process is summarised in Figure 8.
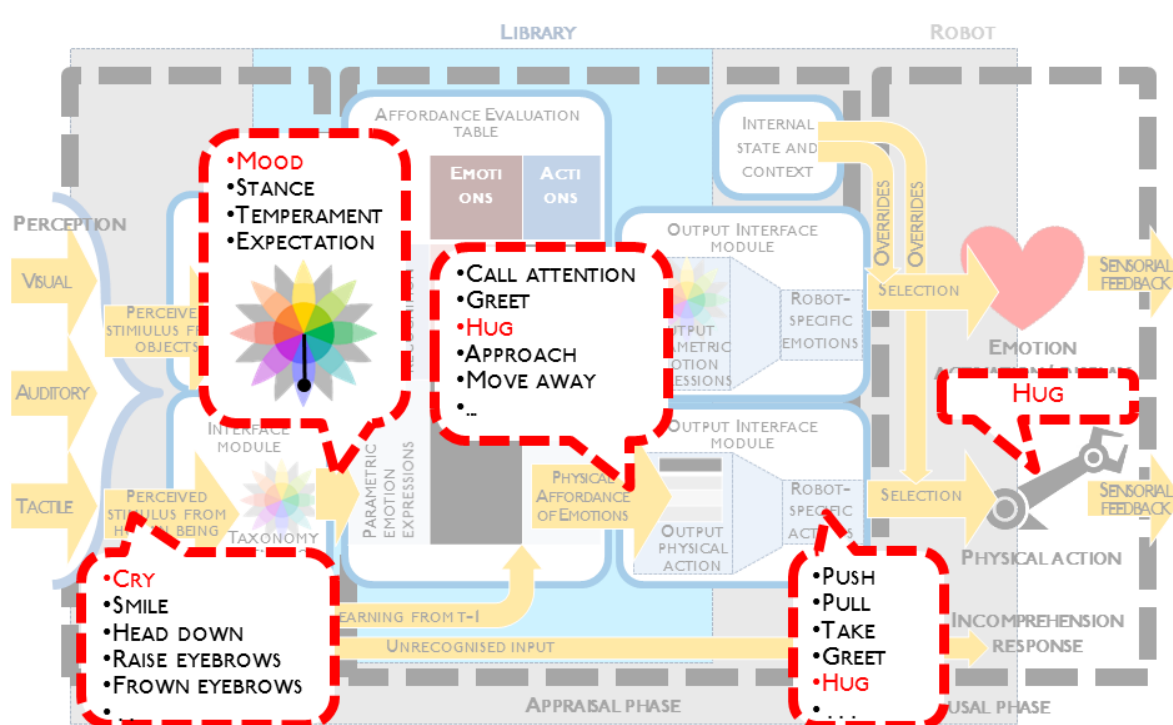

Figure 8. Summary of the object-to-emotion affordance process

4.3. Learning example

Let us suppose that the robot performs the hug suggested by the affordance in the process we have just described. Subsequently, the human reacts. Now at the time T the robot recognises another input, which is an expression of anger.

Before processing further a new affordance, the affordance evaluation table is updated according to the formula in equation 7: $A_{i,j}{}^T = A_{i,j}{}^{T-1} \cdot (1 + \alpha \cdot v)$. Anger has a valence v of -1.

Having defined the learning rate α as 0.2 will result in $A_{i,j}{}^T = 0.9 \cdot (1 + 0.1 \cdot -1) = 0.72$. This update will result in a lower likelihood of the robot to perform the action of hugging in response to a crying human expression.

5. Sample scenarios of application

Our affordances libraries allow us to consider real scenarios in which such mechanisms would help researchers to implement multimodal and functional HRI. For the brevity of this paper we only present two cases which summarise and encapsulate the vast range of possibilities that our approach enables: elderly care and service robot at home.

5.1. Elderly care

In the first of our sample scenarios we consider a relevant problem: (impaired) elderly care by robots [55]–[57]. In this scenario we consider that our libraries could be easily adapted to the specific sensorial or cognitive requests of elderly people: loss of hearing, diminished vision, sensorimotor difficulties, memory losses, loss of expressivity, among others, something that explains why older people report fewer negative emotional experiences and greater emotional control [58]–[60]. Hence,

gerontechnologies [61] must still to be adapted to robotic assistants [62], and the key is to know which affordances thresholds should be modified in order to offer a good service. Consider a possible case: an old man suffering from Alzheimer. He feels uncomfortable and disoriented and the robot observes an abnormal and repetitive pattern of moving, which is understood as a situation of trouble and therefore the robot offers help to the human. The same patient is feeling pain but because of his limited capacity of self-report that assessment relies in large part on observational methods. We will follow affordances in order to allow our robot to evaluate the pain situation [63]:

    i.   Facial expressions: slight frown; sad, frightened face, grimacing, wrinkled forehead, closed or tightened eyes, any distorted expression, rapid blinking.
   ii.   Verbalizations, vocalizations: sighing, moaning, groaning, grunting, chanting, calling out, noisy breathing, asking for help, verbally abusive.
  iii.   Body movements: rigid, tense body posture, guarding, fidgeting, increased pacing, rocking, restricted movement, gait or mobility changes.
  iv.   Changes in interpersonal interactions: aggressive, combative, resisting care, decreased social interactions, socially inappropriate, disruptive withdrawn.
   v.   Changes in activity patterns or routines: refusing food, appetite change, increase in rest periods, sleep, rest pattern changes, sudden cessation of common routines, Increased wandering.
  vi.   Mental status changes: crying or tears, increased confusion, irritability or distress.

Given the detection from video-cameras and voice recording systems, our library could combine data from those specific affordances in order to detect specific requirements previously requested by medical services. All these facial, voice, and body expressions are fundamental clues for the good monitoring of the patient.   In this case, the robot could detect simple pain affordances and offer a response to the patient, or even notice it to the medical personnel. Establishing adequate correlations between medical drugs intake and expected reactions over time, the robot could help to monitorise and even to adapt these prescriptions.

5.2. Service robot at home

The second scenario is related to service robots which could be placed in our homes. From anthropomorphic no non-anthropomorphic robots, these machines will need to interact closely with humans, which are fundamentally emotional [64]. Consider a possible scenario: a robot pet [65]. The user comes back at home being too much tired for playing activities: consequently she is annoyed when the robot tries to play with her. The robot detect this mood and avoids activity although follows the user through the house, as a silent companion. At some point the user addresses to the robot and speaks to it, trying to beg for pardon and asking for direct contact. The robot detects voice tone and combines it with visual clues which show a pleading attitude: therefore the robot approaches to the user and interact with her.
The range of most salient affordances in this scenario would be those related to the robot monitoring of the humans living in that space. The robot must be obviously able to answer to the voice commands of humans, but also should anticipate possible necessities according to the context. Imagine the following scenario: the human user is alone, taking a shower. The robot evaluates the possible danger situation and decides to be there in case of being requested to for any action. Once there, the human asks to the robot for a towel, and it provides a new one to the human. At that point the human looks angry, and the robot evaluates which can be the cause: perhaps it needs a specific kind of towel and some mistake has happened. The robot ask to the human whether the current size of the towel is wrong and whether is needed a bigger or smaller one. The human answers affirmatively: a bigger one. After using it, the human slips and falls. The robot checks the emotional signals: unconscious attitude (closed ayes, without movement), and the same time the bodily analysis shows a lack of movement not related to a contextual space like that. The robot also

monitorises heart rate and skin conductance deciding that the human is in danger and call to the medical services. Meanwhile the ambulance in on the road, the human starts to moan and the robot offers basic touch response to calm him at the same time that use a confident and quiet voice tone to explain the human with the minimal amount of information that a human medical doctor has been called and she will arrive in a feel minutes. At the doctor arrival, the robot provides a basic inform of the process as well as explain the emotional responses of the human at every moment of the report. Both cases show how the monitoring of emotions (or its lack of presence) using face-bodily-tone data and contextual evaluations make possible a better support for the human.

6. Discussion and conclusion

In our implementation of the affordance emotional system, we followed three fundamental requirements:

- Generality:
  we are aiming to make this model widespread and usable by different robots. For this reason, we mainly focus our description on the inputs/outputs and interconnections of the various sub-modules of the architecture, but we do not elaborate too much on the content of each sub-module, that will largely depend on the specific application scenario.
- Low complexity:
  The architecture perform simple operations, because this is supposed to give a fast response that imitates the neural process of affordance perception occurring in the human mind. For this reason, it does not propose complex calculations that would be more appropriate for the reasoning component of a robot's mind.
- Adaptability:
  In affordances, learning and adaptation are fundamental components [4], [66], [67]. Notably, the dynamical nature of our model, which is also multimodal, captures this aspect and allows to adapt the emotional responses of the robot based on the feedback received through repeated interactions with different individual users or groups of users. Indeed, the whole emotional social life is based on emotional interactions, some of them verbal but also non-verbal, a fundamental aspect of HRI that our models covers with detail.

As mentioned earlier, a few emotional affordances models for HRI purposes. Basically, it can be affirmed that all existing models are based on very basic multimodal approaches to affordances. Secondly, these models are non-taking into account contextual and allocentric aspects of emotional human dimensions (including objects, scenarios, and emotions). Thirdly, these models cannot be transferred between different robotic morphologies/designs. Fourth, only our model contemplates a user's dynamical variation across time according a complex mapping of emotional syntax. Finally, our approach is more complete, using a broad range of emotional variations (Plutchik's model), beyond the classic approaches using the basic Ekman's 6 emotional values; this makes possible obtain a more accurate monitoring, understanding and evaluation of emotional affordances for HRI purposes.

Future work. We plan to obtain a more accurate monitoring, understanding and evaluation of emotional affordances for HRI purposes. Our intention is to use the libraries in different morphological robotic systems, checking the adaptability of our libraries to the data integration made using several general designs.

Supplementary Materials: None

Author Contributions: G.T., J.V., and L.J. conceived and designed the research idea; G.T. carried out the implementation; L.J. and J.V. wrote the theoretical backgrounds of the paper.

Conflicts of Interest: The authors declare no conflict of interest. The founding sponsors had no role in the design of the study; in the collection, analyses, or interpretation of data; in the writing of the manuscript, and in the decision to publish the results.

References

[1]   J. J. Gibson, *The senses considered as perceptual systems*, 1st ed. Boston: Houghton. Mifflin Company, 1966.

[2]   J. J. (James J. Gibson, *The ecological approach to visual perception*. Lawrence Erlbaum associates, 1979.

[3]   J. J. Gibson, "The concept of affordances," in *Perceiving, Acting, and Knowing: Toward an Ecological Psychology*, no. December 2007, 1977, pp. 67–82.

[4]   E. Bardone, "Affordances as abductive anchors," in *Studies in Computational Intelligence*, 2010, vol. 314, pp. 135–157.

[5]   E. J. Gibson, "Where Is the Information for Affordances?," *Ecol. Psychol.*, vol. 12, no. 1, pp. 53–56, 2000.

[6]   L. Jamone, E. Ugur, A. Cangelosi, L. Fadiga, A. Bernardino, J. Piater, and J. Santos-Victor, "Affordances in Psychology, Neuroscience, and Robotics: A Survey," *IEEE Trans. Cogn. Dev. Syst.*, vol. 10, no. 1, pp. 4–25, 2018.

[7]   A. R. Damasio, "Descartes' error: Emotion, rationality and the human brain," *New York Putnam*, p. 352, 1994.

[8]   A. R. Damasio, *The Feeling of What Happens: Body and Emotion in the Making of Consciousness*, vol. 401, no. 6756. 1999.

[9]   M. Minsky, "The emotion machine," *New York Pantheon*, pp. 1–7, 2006.

[10]   J. Vallverdú and D. Casacuberta, "Handbook of Research on Synthetic Emotions and Sociable Robotics: New Applications in Affective Computing and Artificial Intelligence (Google eBook)," *Handb. Res. Synth. Emot. Sociable Robot. New Appl. Affect. Comput. Artif. Intell.*, pp. xxiii–xxvi, 2009.

[11]   J. Vallverdú, M. Mazzara, M. Talanov, S. Distefano, and R. Lowe, *Advanced Research on Biologically Inspired Cognitive Architectures*. 2017.

[12]   J. Vallverdu, *Handbook of research on synthesizing human emotion in intelligent systems and robotics*. 2015.

[13]   J. Vallverdú and G. Trovato, "Emotional affordances for human-robot interaction," *Adapt. Behav.*, vol. 24, no. 5, pp. 320–334, Oct. 2016.

[14]   V. Franzoni, A. Milani, and J. Vallverdú, "Emotional affordances in human-machine interactive planning and negotiation," in *Proceedings - 2017 IEEE/WIC/ACM International Conference on Web Intelligence, WI 2017*, 2017.

[15]   D. Matsumoto, "Paul Ekman and the legacy of universals," *J. Res. Pers.*, vol. 38, no. 1, pp. 45–51, 2004.

[16]   P. Ekman, "An argument for basic emotions," *Cognition & Emotion*, vol. 6, no. 3. pp. 169–200, 1992.

[17]   J. Vallverdu, "Ekman's Paradox and a Naturalistic Strategy to Escape From It," *Int. J. Synth. Emot.*, vol. 4, no. 2, pp. 1–7.

[18]   M. J. Schroeder and J. Vallverdú, "Situated phenomenology and biological systems: Eastern and Western synthesis," *Progress in Biophysics and Molecular Biology*, vol. 119, no. 3. pp. 530–537, 2015.

[19]   J. Prinz, "Embodied emotions," *Think. about Feel.*, pp. 44–58, 2004.

[20]   P. M. Niedenthal and M. Maringer, "Embodied Emotion Considered," *Emot. Rev.*, vol. 1, no. 2, pp. 122–128, 2009.

[21]   P. M. Niedenthal, "Embodying emotion," *Science*, vol. 316, no. 5827. pp. 1002–1005, 2007.

[22]   P. M. Niedenthal and M. Brauer, "Social Functionality of Human Emotion," *Annu. Rev. Psychol.*, vol. 63, no. 1, pp. 259–285, 2012.

[23]   D. T. Neal and T. L. Chartrand, "Embodied emotion perception: Amplifying and dampening facial feedback modulates emotion perception accuracy," *Social Psychological and Personality Science*, vol. 2, no. 6. pp. 673–678, 2011.

[24]   C. Darwin, "The expression of the emotions in man and animals," *London, UK John Marry*, p. 374, 1872.

[25]   L. Nummenmaa, E. Glerean, R. Hari, and J. K. Hietanen, "Bodily maps of emotions.," *Proc. Natl. Acad. Sci. U. S. A.*, vol. 111, no. 2, pp. 646–51, 2014.

[26]   C. Nicolaidis, C. C. Kripke, and D. Raymaker, "Primary care for adults on the autism spectrum," *Medical Clinics of North America*, vol. 98, no. 5. pp. 1169–1191, 2014.

[27]   C. Nicolaidis, D. Raymaker, K. McDonald, S. Dern, W. C. Boisclair, E. Ashkenazy, and A. Baggs, "Comparison of healthcare experiences in autistic and non-autistic adults: A cross-sectional online survey facilitated by an academic-community partnership," *J. Gen. Intern. Med.*, vol. 28, no. 6, pp. 761–769, 2013.

[28]   M. Iacoboni, "Imitation, Empathy, and Mirror Neurons," *Annu. Rev. Psychol.*, vol. 60, no. 1, pp. 653–670, 2009.

[29]   M. Iacoboni, "Understanding others: imitation, language, empathy," *Perspect. Imitation From Cogn. Neurosci. to Soc. Sci.*, pp. 56–77, 2005.

[30]   A. Chemero, *Radical embodied cognitive science*. Cambridge (MA): MIT Press, 2009.

[31]   P. Vicente, L. Jamone, and A. Bernardino, "Online Body Schema Adaptation Based on Internal Mental Simulation and Multisensory Feedback," *Front. Robot. AI*, vol. 3, 2016.

[32]   J. LeDoux, *The Emotional Brain*, vol. 256. 1996.

[33]   L. JE, "Sensory systems and emotions: A model of affective processing," *Integr Psychiatry*, vol. 4, p. 237, 1986.

[34]   J. Bryson and E. Tanguy, "Simplifying the Design of Human-Like Behaviour," *Int. J. Synth. Emot.*, vol. 1, no. 1, pp. 30–50, 2010.

[35]   D. A. Norman, *Emotional design*. 2005.

[36]   H. Osawa, J. Mukai, and M. Imai, "Anthropomorphization of an object by displaying robot," in *Proceedings - IEEE International Workshop on Robot and Human Interactive Communication*, 2006, pp. 763–768.

[37]   H. Osawa, Y. Matsuda, R. Ohmura, and M. Imai, "Embodiment of an agent by anthropomorphization of a common object," *Web Intell. Agent Syst.*, vol. 10, no. 3, pp. 345–358, 2012.

[38]   A. Richert, S. Müller, S. Schröder, and S. Jeschke, "Anthropomorphism in social robotics: empirical results on human–robot interaction in hybrid production workplaces," *AI and Society*, pp. 1–12, 2017.

[39]   P. Ekman, "Basic Emotions," in *Handbook of cognition and emotion*, 1999, pp. 45–60.

[40]   T. Bänziger, D. Grandjean, and K. R. Scherer, "Emotion Recognition From Expressions in Face, Voice, and Body: The Multimodal Emotion Recognition Test (MERT)," *Emotion*, vol. 9, no. 5, pp. 691–704, 2009.

[41]   M. Talanov, J. Vallverdú, S. Distefano, M. Mazzara, and R. Delhibabu, "Neuromodulating cognitive architecture: Towards biomimetic emotional AI," in *Proceedings - International Conference on Advanced Information Networking and Applications, AINA*, 2015, vol. 2015–April, pp. 587–592.

[42]   R. E. Nisbet, *The Geography of Thought: How Asians and Westerners Think Differently...and Why: Richard E. Nisbett: 9780743255356: Amazon.com: Books*. New York: Free Press (Simon & Schuster, Inc.), 2003.

[43]   A. Norenzayan and R. E. Nisbett, "Culture and causal cognition," *Curr. Dir. Psychol. Sci.*, vol. 9, no. 4, pp. 132–135, 2000.

[44]   J. Vallverdú, T. Nishida, Y. Ohmoto, S. Moran, and S. Lázare, "Fake empathy and human- robot interaction (HRI): A preliminary study," *Int. J. Technol. Hum. Interact.*, vol. 14, no. 1, 2018.

[45]   R. Plutchik, "Psychoevolutionary Theory of Basic Emotions," *Am. Sci.*, no. February, p. 2007, 1980.

[46]   J. Sun, J. L. Moore, A. Bobick, and J. M. Rehg, "Learning visual object categories for robot affordance prediction," *Int. J. Rob. Res.*, vol. 29, no. 2–3, pp. 174–197, 2010.

[47]   Y. Sun, S. Ren, and Y. Lin, "Object-object interaction affordance learning," *Rob. Auton. Syst.*, vol. 62, no.

4, pp. 487–496, 2014.

[48]　G. Trovato, M. Zecca, T. Kishi, N. Endo, K. Hashimoto, And A. Takanishi, "Generation Of Humanoid Robot's Facial Expressions For Context-Aware Communication," *Int. J. Humanoid Robot.*, Vol. 10, No. 01, P. 1350013, 2013.

[49]　H. Miwa, T. Okuschi, H. Takanobu, A. Takanishi, and H. Miwal, "Development of a New Human-like Head Robot WE-4," *Proc. 2002 IEEE/RSJ Intl. Conf. Intell. Robot. Syst.*, no. October, pp. 2443–2448, 2002.

[50]　J. Decety, "The neurodevelopment of empathy in humans," *Dev. Neurosci.*, vol. 32, no. 4, pp. 257–267, 2010.

[51]　R. Debes, "Empathy and mirror neurons," in *The Routledge Handbook of Philosophy of Empathy*, 2017, pp. 54–63.

[52]　J. Vallverdú and D. Casacuberta, "Ethical and technical aspects of emotions to create empathy in medical machines," *Intell. Syst. Control Autom. Sci. Eng.*, vol. 74, pp. 341–362, 2015.

[53]　I. Poggi and C. Pelachaud, "Performative faces," *Speech Commun.*, vol. 26, no. 1–2, pp. 5–21, 1998.

[54]　I. Poggi, *Le parole del corpo. Introduzione alla comunicazione multumodale*. Carocci, 2006.

[55]　J. Broekens, M. Heerink, and H. Rosendal, "Assistive social robots in elderly care: a review," *Gerontechnology*, vol. 8, no. 2, pp. 94–103, 2009.

[56]　C. Leroux, O. Lebec, M. W. Ben Ghezala, Y. Mezouar, L. Devillers, C. Chastagnol, J. C. Martin, V. Leynaert, and C. Fattal, "ARMEN: Assistive robotics to maintain elderly people in natural environment," *IRBM*, vol. 34, no. 2, pp. 101–107, 2013.

[57]　Y. Wu, M. Chetouani, V. Cristancho-Lacroix, J. Le, and A. Rigaud, "ROBADOM: An Assistive Robot for the Elderly with Mild Cognitive Impairment: System design and Users' Perspectives," in *5th CRI (Companion Robotics Institute) Workshop AAL User-Centric Companion Robotics Experimentoria, Supporting Socio-ethically Intelligent Assistive Technologies Adoption*, 2011, pp. 1–12.

[58]　J. J. Gross, L. L. Carstensen, M. Pasupathi, J. Tsai, C. G. Skorpen, and A. Y. C. Hsu, "Emotion and aging: Experience, expression, and control," *Psychol. Aging*, vol. 12, no. 4, pp. 590–599, 1997.

[59]　M. Altamura, F. A. Padalino, E. Stella, A. Balzotti, A. Bellomo, R. Palumbo, A. Di Domenico, N. Mammarella, and B. Fairfield, "Facial Emotion Recognition in Bipolar Disorder and Healthy Aging.," *J. Nerv. Ment. Dis.*, vol. 204, no. 3, pp. 188–193, 2016.

[60]　K. W. Burton and A. W. Kaszniak, "Emotional experience and facial expression in Alzheimer's disease.," *Aging, Neuropsychol. Cogn.*, vol. 13, no. April 2013, pp. 636–51, 2006.

[61]　P. Pires, L. Mendes, J. Mendes, R. Rodrigues, and A. Pereira, "Integrated e-Healthcare System for Elderly Support," *Cognit. Comput.*, vol. 8, no. 2, pp. 368–384, 2016.

[62]　T. Shibata and J. F. Coughlin, "Trends of robot therapy with neurological therapeutic seal robot, PARO," *Journal of Robotics and Mechatronics*, vol. 26, no. 4. pp. 418–425, 2014.

[63]　P. on C. P. in O. P.　American Geriatrics Society, "The management of chronic pain in older persons.," *J. Am. Geriatr. Soc.*, vol. 46, no. 5, pp. 635–651, 1998.

[64]　A. Bonarini, "Can my robotic home cleaner be happy? Issues about emotional expression in non-bio-inspired robots," *Adapt. Behav.*, vol. 24, no. 5, pp. 335–349, 2016.

[65]　W. P. Lee, J. W. Kuo, and P. C. Lai, "Developing emotion-based pet robots," in *Lecture Notes in Electrical Engineering*, 2009, vol. 39 LNEE, pp. 389–400.

[66]　N. Valanides, "Technological tools: From technical affordances to educational affordances," *Problems of Education in the 21st Century*. 2018.

[67]　K. Cheng and C. Tsai, "Affordances of augmented reality in science learning: suggestions for future research," *J. Sci. Educ. Technol.*, vol. 22, no. 4, pp. 449–462, 2013.

[68]   S. Ambroszkiewicz, W. Bartyna, M. Szymczakowski, and K. Skarżyński, "Affordances and SOA-based multi-robot systems," *Adv. Intell. Syst. Comput.*, vol. 323, pp. 177–187, 2015.

[69]   F. Cid, "Learning Emotional Affordances based on Affective Elements in Human-Robot Interaction Scenarios," *Waf2014*, no. June, pp. 1–10, 2014.

[70]   C. Wang, K. V. Hindriks, and R. Babuska, "Robot learning and use of affordances in goal-directed tasks," in *IEEE International Conference on Intelligent Robots and Systems*, 2013, pp. 2288–2294.