Tangible Interaction and Embodied Cognition challenged by remote control issues

Vincent Ferrari[1], Valentin Braud[1,2,3], Laurent Bovet[2,3] and Nadine Couture[3]

[1]CREA, Aix-Marseille université, École de l'air et de l'espace, F-13661 Salon-de-Provence, France
[2]ELISA Aerospace, F-02100 Saint-Quentin, France
[3]Univ. Bordeaux, ESTIA INSTITUTE OF TECHNOLOGY, F-64210 Bidart, France

Abstract
This article shows the interest of considering embodied cognition as a theoretical framework for tangible interfaces. This leads to use tangible interfaces as technical solutions to enable the enforcement of embodied cognition principles. The reflection that was carried out in this article is motivated by the context of remote UAV systems control.

Keywords
Embodied Cognition, Tangible Interaction, Unmanned Aerial System Control, Interdisciplinary

1. Introduction

An efficient control of a remote system (drone, rovers, etc.), necessarily implies the development of the best possible spatial representation in the operator's mind. In concrete terms, what does a drone operator have at his disposal to best represent an aeronautical situation that sometimes takes place thousands of kilometres away? First of all, the operator processes the information provided by the ground station interfaces. The quality of these interfaces (the information presented) will determine the plausibility of the operator's spatial representation. Furthermore, to ensure this plausibility, the experienced operator also has his/her previous knowledge about unmanned aircraft vehicles. Moreover, the operator's level of experience will modulate the relevance of their spatial representation. Finally, it is necessary to ensure that the information provided by the interfaces is in line with the nature of the operators' knowledge. This being said, do the interfaces of UAS ground stations currently meet these requirements? The answer is no. We argue here that, in order to address the needs of UAS remote control issues, particularly those relating to the quality of the operator's spatial representation, alternatives to traditional graphical user interfaces (GUIs) implemented in ground stations must be provided. From our perspective, with the support of the embodied cognition theoretical framework, tangible interaction will better exploit the capabilities of operators in the remote control of UAS. The two disciplines addressed here, namely embodied cognition theory (shortened to EC for future reference) and tangible interaction (shortened to TI for future reference), will challenge each other. The two disciplines must cohabit and contribute to each other in order to respond to

CEUR Workshop Proceedings (CEUR-WS.org)

the very strong constraints of the application field that carries very high stakes, namely drone systems, and more generally systems that are controlled remotely.

In this particular context, the user (the sensor operator or the pilot) must manage a very large amount of information enabling him to reconstruct a "reality" that is distant. So far, this reconstruction has been performed through screens which are unfortunately multiple and require a purely cognitive processing of the information, and thus rely on the individual's information processing capabilities and on the visual modality. It is therefore very interesting to consider an alternative to this full visual and cognitive approach, thanks to TI sustained by EC. A gradual shift in the cognitive sciences towards embodied paradigms of human cognition has inspired us to think about how EC should theoretically support TI.

2. Embodied Cognition

EC is a vision developed in opposition to traditional cognitivism. EC, which is still a minority today, is opposed to the analogy between the functioning of the human brain and the functioning of a computer. This analogy had been the foundation of cognitivism. The climax of the disagreement between cognitivism and EC lies in the format of knowledge representation. According to EC, the mind cannot be reduced to the amodal processing of symbolic information. On the contrary, cognition would be rooted in sensory and motor systems. Cognition would therefore no longer be abstract and amodal, but rather essentially sensorimotor. In short, EC considers that the mind must be understood in the context of its body (the 'sensorimotor context'), and of its interaction with the environment [1]. Therefore, human knowledge should be sensorimotor in nature. Thus, if the cognitive system is by essence sensorimotor, then the spatial representations, elaborated by the human brain to guide actions, must also integrate sensory and motor information. It should be noted here that the notion of 'spatial representation' refers to the way in which an individual represents the elements of an environment as well as their absolute and relative positions [2]. According to the theoretical principles of EC, a spatial representation is constructed from two sources: an external source, perception, and an internal source, memory. If the information at the source of spatial representations is embodied (in the sense that it is rooted in the body), then it is reasonable to think that spatial representations are also embodied [3].

To summarise, EC is a model of cognition built on the assumption that t's not possible to separate the elements of the triptych: action, perception and environment [4].

3. Could Embodied Cognition be a theoretical framework for Tangible Interaction design?

A theoretical framework introduces and describes the theory or theories underpinning a research problem. Thus, theoretical frameworks support research by describing and/or drawing from relevant theoretical aspects obtained in previous work. We argue that EC should be, in the future, a theoretical framework for TI. We are at an exploratory level, but we believe that EC can provide the theoretical elements that would allow interaction designers to better understand and therefore anticipate what may or may not work if they use a system with tangible interfaces.

Therefore, amongst the concepts we have identified that link EC & TI, we will discuss below only two of us which appear to us as fundamental: affordance and legibility.

3.1. Affordance

Affordance in TI. For Norman [5], affordance is the characteristic of an object to 'naturally' suggest to its user how to use it. Affordant props allow intuitive and rich manipulations. This is because the manipulation of props in the tangible system relies on the perception the user has of his/her environment, and exploits the dexterity of the user. Thus, affordant props must both enable a digital effect analogous to a real effect (metaphor of verb in [6]) and, through its physical design, suggest to the user all the possibilities of action (for instance, a cup handle invites us to grasp it) (metaphor of noun in [6]). Indeed, to determine whether an interface is tangible, Fishkin in [6] introduced these two kinds of metaphor. The metaphor of the verb in the sense of action helps to answer the question: how does the effect of a user's action on props compare to the effect of that same action in the real world? If the action performed by the user within the tangible system is a faithful reproduction of the action that the user would perform in the real world, then the system can be said to offer a metaphor of verb. The designer of a tangible interface also chooses the shape, colour, smell, weight and texture of its props. In this way, he generates sensory links between the user and the manipulated props, and thus, between the user and the data embodied by the props. This is the metaphor of noun [1] . The system will be considered offering a metaphor of nouns if the appearance of the object strongly looks like something that the user knows. A tangible interface can propose both a metaphor of verb and a metaphor of noun. In this case, it makes strong use of analogies of both appearance and use, but the physical and virtual objects remain different. Finally, if more analogies are not needed, the metaphor is said to be full, the virtual system and the real system do only one.

What EC says about affordance? While GUIs rely less, or not at all, on the sensory-motor abilities of operators, tangible user interfaces by their physical existence enable a more intuitive integration of the digital world into the users' environment. In other words, the 'tangibility' of the props makes it possible to provide affordances that will be perceptible by the human sensory-motor system. The fact that the props can be manipulated makes it easier to convoke the user's proprioceptive abilities, which is impossible with graphical interfaces. Phillips & Ward [7] argue that humans perceive the world in the form of affordances, i.e. that the environment incorporates cues that univocally suggest actions to humans. This hypothesis assumes the existence of a rapid and systematic link between perception and action that optimizes the decision making process in the environment. In other words, affordances are visual cues that incorporate information about action possibilities in the environment and are automatically 'caught' by the human sensorimotor system. Taking, the same example as previously, cup handle placed near a person is an affordance, because the handle calls for the person to grasp it. Note that, if this cup is outside the individual's zone of possible action, the motor cortex of the person will not be activated. The possibility of action is therefore the boundary between peripersonal space (where action is possible directly) and extrapersonal space (where action is possible as a result of movement) [8].

---

[1]The authors of [6] chose these terminologies 'metaphor of noun' and 'metaphor of verb' because cognitive psychologists claim that nouns and verbs seem to be deeply embedded in our consciousness.

To conclude about affordance by mixing EC and TI, we can say that, for EC, perception does not only have the passive role of representing information from the environment (as opposed to traditional cognitivism models), but perception guides action through the processing of affordances. While GUIs rely less, or not at all, on the sensorimotor abilities of operators, tangible user interfaces by their physical existence allow for a more intuitive integration of the digital world into the users' environment. In other words, the 'tangibility' of the props makes it possible to provide affordances that will be perceptible by the sensorimotor system. The fact that the props can be manipulated makes it easier to use the user's proprioceptive abilities, which is impossible with graphical interfaces. To summarise, and in accordance with an embodied vision of cognition, perception would have the function of guiding action.

3.2. Legibility

Legibility in TI. In 1992, Durrell Bishop designed the marble answering machine, described in [9]. It is the first example of an attempt to link the physical world to the digital world. The concept is a telephone answering machine with tangible balls[2]. This system marks the beginning of digital design. Bishop emphasises that, in the interest of communication, what is important is not just what objects do, but what they convey, and how they convey it. He considers that there is "*an immediate legibility* of the object that contrasts with *the illegibility of computing*. This remark is of course strongly related to the notion of affordance presented above. We consider legibility in the sens given by D. Bishop and as part of the LAVA heuristics (Legible, Actionable, Veritable, and Aspirational). To go further on the different facets of legibility see the 7 pages, section A.2.1, of the appendices of [10].

What does EC say about legibility? According to the principles of EC, perception encodes environmental information according to the possibilities of action, i.e. according to the different senses and motor actions solicited by the environment (sometimes in the form of affordances). Considering that perception is based on the sensorimotor system, it seems logical that memory is also based on this system, since one of its functions is to store the information provided by perception (i.e. sensorimotor information). It is the procedural memory (a subdivision of human memory) which groups together the perceptual, motor and cognitive representations stored in long-term memory and likely to be processed in working memory. These are dynamic representations that allow the acquisition and realisation of various sensorimotor skills. Access to this memory is automatic (particularly in the presence of affordances). According to Tucker and Ellis [11], the perception of an object that has already been manipulated reactivates the possibilities of actions on that object. When an object is perceived, direct processing is done. Glenberg [12] considers that affordances are not only perceptible information in the environment through possible actions, but are complemented by the observer's prior knowledge of the object's functions or experiences with the object which leads to this legibility.

To conclude about legibility by mixing EC and TI, we can say that, unlike graphical interfaces, which are overloaded with digital information that requires costly 'translation' to move from a symbol to an action, a tangible interface presents digital data in the form of manipulable

---

[2]The machine physically displays voice messages with beads, which can then be listened to, deleted in any order.To listen to a message, the user takes the ball and places it on a defined slot on the machine. Then the message can be deleted or the user can also choose to store the messages, outside the machine in a container

and affordant objects to optimise processing. In other words, tangible interfaces modify the nature of the information to be processed so that it 'matches' the sensorimotor nature of the knowledge stored in the operator's memory. For the human brain, symbols presented on a screen that represent the state of a system are not directly legible (they require 2 translations, one from symbol to understanding and another to action)! The embodied approach to cognition reduces this distance by activating more physical modes of processing, more primary ones, more adaptive, and therefore faster modes of processing. Exactly like in TI which makes digital data legible for humans by "speaking the language of embodiment"!

4. Conclusion and Perspective

The previous discussion on affordance and on legibility, is a first step to be continued, showing EC as a relevant theoretical framework for TI design. Moreover, TI appears as perfect technical solution to enable the enforcement of EC principles. This discussion was motivated by the context of UAS controls where current GUIs have been designed to provide as much information as possible to operators. With missions diversification, the complexity and the number of information have increased. As they reported to us in informal meetings, it is presently hard for operators to find needed information easily and rapidly. Taking into account the fact that most operators are former pilots or at least trained pilots, by integrating props into the interface, it is possible to improve the operator's ability to understand the system and its operation. Thus, by improving the way the information is understood and manipulated by the operators, the difficulties linked to the interface are reduced, allowing them to focus more on the operational elements (mission objectives, payload status, communication, etc.).

The originality of this operational environment is that the digital data within the system comes from a physical reality! Indeed, it is a remote control via digital data. There are many other fields of application that would benefit from the possibility of embodying digital data that translate a physical reality. e.g. the remote control inside a nuclear power plant, but also space and aeronautics, i.e. dirty and dangerous environments. Beyond the claimed poetic and metaphorical aspects of TI, beyond the aestheticism and the often artistic side of TI, we are convinced, and this has been discussed in [13], that TI is also very useful! For some critical systems, TI combined with more classical interfaces can improve the systems:maintaining one's drone system, not dropping it on someone, monitoring a nuclear power plant, or planning a space walk. TI will offer the operator not only symbolic data from classic graphical interfaces but also the meaning of it. These domains definitely need this embodiment and a theoretical framework to anticipate the effects and limits of this embodiment, and we believe that this framework can be provided by the theory of EC.

This leads us to ask the following question: can we still talk about 'tangible interaction' when we embody digital data that has a physical reality? This question appears to us as a real intellectual and theoretical challenge. It is a "mise en abyme": giving a physical reality to virtual data that represents a physical reality! This introduces the notion of embodiment of digital data that represents a physical reality that is not accessible because it is remote.

References

[1] L. W. Barsalou, Perceptual symbol systems, Behavioral and Brain Sciences 22 (1999) 577–660. doi:10.1017/S0140525X99002149.

[2] R. A. Zwaan, G. A. Radvansky, Situation Models in Language Comprehension and Memory (1998) 24. doi:10.1037/0033-2909.123.2.162.

[3] L. Dutriaux, V. Gyselinck, Cognition incarnée : un point de vue sur les représentations spatiales:, L'Année psychologique Vol. 116 (2016) 419–465. doi:10.3917/anpsy.163.0419.

[4] L. W. Barsalou, Grounded Cognition, Annual Review of Psychology 59 (2008) 617–645. doi:10.1146/annurev.psych.59.103006.093639.

[5] D. A. Norman, Affordance, conventions, and design, Interactions 6 (1999) 38–43. doi:10.1145/301153.301168.

[6] K. Fishkin, A taxonomy for and analysis of tangible interfaces, Personal and Ubiquitous Computing 8 (2004). doi:10.1007/s00779-004-0297-4.

[7] J. C. Phillips, R. Ward, S-R correspondence effects of irrelevant visual affordance: Time course and specificity of response activation, Visual Cognition 9 (2002) 540–558. doi:10.1080/13506280143000575.

[8] E. Y. Coello, A. Bartolo, Language and Action in Cognitive Neuroscience (2013) 407. doi:10.4324/9780203095508.

[9] B. Moggridge, Designing Interactions, MIT Press, Cambridge, Mass, 2007.

[10] B. Ullmer, O. Shaer, A. Mazalek, C. Hummels, Weaving Fire into Form: Aspirations for Tangible and Embodied Interaction, volume 44, 1 ed., Association for Computing Machinery, New York, NY, USA, 2022.

[11] M. Tucker, R. Ellis, On the Relations Between Seen Objects and Components of Potential Actions (1998) 17. doi:10.1037/0096-1523.24.3.830.

[12] A. M. Glenberg, What memory is for, Behavioral and Brain Sciences 20 (1997) 1–19. doi:10.1017/S0140525X97000010.

[13] P. R. Cohen, D. R. McGee, Tangible multimodal interfaces for safety-critical applications, Communications of the ACM 47 (2004) 41–46. doi:10.1145/962081.962103.