Representation-phobia and the complexity of embodied interaction

Thomas Markussen, Assistant Professor, PhD, MA
Institute for Design, Aarhus School of Architecture, Denmark

Abstract

In current interaction design research there is a widespread belief that situated action and embodied interaction should replace mental representations in the theoretical account of human cognition. This exclusion of representation is however diagnosed as a sign of representation-phobia by Anderson (2003) who claims that it is misguided. This paper aims to show why and how it can be overcome. Initially, a literature review will show how representation-phobia manifests itself through two different versions in HCI research. On the basis of this I argue that representation-phobia leads to a theoretical dead end. Then, by drawing on semiotics and recent findings from cognitive research, I argue that we cannot understand the rich complexity of embodied interaction unless we furnish our thinking with a dynamic notion of representation.

Keywords: cognition, embodied interaction, representation, experiential knowledge, semiotics

In interaction design research, the two notions of "embodied interaction" and "situated action" are thought to be useful for understanding how to design interactive artifacts or environments that resonate with the rich complexity of our bodily and everyday interaction with the material world (Dourish, 2004; see e.g. Hornecker, 2005; Klemmer, Hartmann, & Takayama, 2006).

Despite their divergent foci of interests, proponents of embodied interaction and situated action generally share the view that cognition is not an internal affair in the head of the user, but something that is shaped and evolves through our actions as we react to real-time requirements from mundane settings. Further, in much recent work, the increased focus on action and bodily experience is coupled with skepticism towards the role of mental representations (Rowlands, 2009, p. 127). In fact, the adjectives "embodied" and "situated" are sometimes used as synonyms for the idea that cognition unfolds in a direct manner without the intervention of mental representations and that HCI would therefore be better off if is simply deleted the concept of representation from its terminological vocabulary. For this reason, embodied interaction and activity theorists often see their frameworks as being diametrically opposed to traditional HCI and classical cognitivism.

However, as Anderson has warned us, the just critique of traditional HCI's notion of representation should not lead to a "representation-phobia" in interaction design research. Thus, according to Anderson, it is a mistake to exclude representations from frameworks of embodied interaction. What embodied interaction is essentially about is a re-thinking, rather than a rejection of representations (Hutchins, 2005; Rowlands, 2006, 2009; Sinha, 2005; Zlatev, 2005)

Interestingly, a similar warning is echoed in Suchman (2007) who is careful to underline, in the recently republished edition of *Plans and situated action*, that the supplanting of mental representations with action is a theoretical dead end. True, the nature and foundation of mental representations must be re-conceptualized, but embodied and situated cognition relies essentially on a subtle interweaving between action and mental representation.

The aim of this paper is above all to cure interaction design theory of its representation-phobia by showing that, contrary to what is being claimed, mental representations do in fact play a foundational role in user experience – even at the most basic levels of tangible and physical interaction. To set the scene, I begin with a literature review focusing on two different forms of representation-phobia and the common cause of their outbreak. First there is representation-phobia as it comes to the fore in ideas of embodied interaction claiming that our abstract reasoning and making sense of the world is primarily made up from meaning structures the nature of which is physical and spatial, not symbolic and representational. Second, there is representation-phobia as found in activity theory where it is reflected in the belief that mental representations are too static and inflexible for coping with the ever changing and ill-structured nature of everyday situations (for exceptions see Bærentsen & Trettvik, 2002).

Having identified the key characteristics of representation-phobia, I will draw on semiotics and findings from recent cognitive research in order to show why it must be abandoned. More specifically, I will argue that the discovery of canonical neurons show that representations are present at the most basic level of embodied interaction. This presupposes of course a well-defined notion of what a representation is. Therefore, I will provide such a notion together with a stratified model showing how low-level representations grow into more sophisticated symbolic forms. To improve the understandability of my argument, I will provide as many examples as possible throughout the paper.

1. Literature review: Representation-phobia in theories of embodied interaction and situated action

Since its early conception in the 1950s, research into the design of Human-Computer Interaction (HCI) has been a cross-disciplinary scientific enterprise encompassing computer science, cognitive psychology and Artificial Intelligence (AI). Out of this original core of HCI-disciplines various theories of the human mind and cognition have been developed over the years in order to understand how to design human-computer interaction in harmony with principles of human thinking and experience (see e.g. Carroll, 1987, 1991).

As is well known, theories of situated cognition and embodied interaction have been developed because of an increasing dissatisfaction with the once so dominant symbol system approach to cognition. According to this approach, which was originally founded by Newell and Simon (Newell, 1980; Newell & Simon, 1972) cognition basically consists in the algorithmic manipulation of symbol structures. Through the work of Newell and Simon's successors these symbol structures have become known in HCI literature as goal-oriented plans, scripts (1977) frames (1975) and mental models (Norman, 1988; Payne, 2003). Generally speaking, these concepts refer to the idea that we use stereotypical background knowledge as structures for comprehending events and objects reoccurring in everyday life. The basic assumption is that these knowledge representations are built in as part of a prior knowledge base in long-term memory, which we can recruit from for the purpose of planning and executing action relevant in a given situation. For instance, according to Schank & Abelson (1977) we presumably possess a RESTAURANT script representing background knowledge for a sequence of actions relevant to perform when visiting a such a place, namely entering >> get seated >> ordering >> eating >> paying >> exiting.

There are two main objections that have been raised against the notion of such inner symbols. First, it is claimed that symbolic knowledge representations are unable to account for how we handle problems and execute actions in everyday settings. Because representations, under the symbolic view, are algorithmic and presuppose pre-existing and static knowledge structures, while everyday situations, so the argument goes, are unbound and ill structured and therefore not compatible with this format (Winograd & Flores, 1986). Central for the consolidation of this objection was the work of Suchman (1987) who convincingly argued for the role played by the resources of the immediate situation in shaping human action. People are improvisatory and employ ad hoc rules for reasoning instead of abstract algorithmic rules for decision-making and inferences (Kaptelinin & Nardi, 2006, p. 16).

The second objection is that the symbol system approach completely ignores the role of the body for our understanding and making sense of the world. According to Newell and Simon inner symbols are "substrate neutral" in the sense that the nature of our bodies and perceptual apparatus is regarded as irrelevant for the constitution and structuring of their content. This is evidenced by a passage in *Human Problem Solving*, where Newell and Simon overly admit that they have omitted both low-level "sensory and motor skills, and many aspects of perception" from their study of symbols, because they do not consider these factors central for the constitution of symbolic activity (Newell & Simon, 1972, p. 8). However, as the proponents of the embodied mind paradigm have convincingly argued, we cannot understand human thinking and reasoning unless we take structures derived from our perceptual and bodily interaction with the physical world into account (Johnson, 1987; Lakoff, 1987; Lakoff & Johnson, 1999; Varela, Thompson, & Rosch, 1991).

Situated Action and Embodied Interaction are two conceptual frameworks that were introduced in HCI during the late 1980s and 1990s in order to compensate for the theoretical shortcomings of traditional HCI. While situated action is essentially about trying to understand how to design technological systems based on the principles of our online cognitive activity in the world (see e.g. Kaptelinin & Nardi, 2006; Nardi, 1996; Suchman, 1987; Wilson & Clark, 2005; Winograd & Flores, 1986), embodied interaction is a design theory focusing on how to design intuitive, tangible, graspable and direct physical interaction with technology (Dourish, 2004; Ehn & Linde, 2004; Hornecker, 2005; Hurtienne & Israel, 2007; Lund, 2003; Lund & Waterworth, 1998).

What these two approaches have in common, despite their many profound differences and divergent foci of interests, is a widespread distrust of the concept of internal symbols or any sign of mental knowledge representations. In fact, the adjectives of "situated" and "embodied" are often meant to put emphasis on the assumption that cognition unfolds in a *direct and distributed* manner and that human-computer interaction should therefore be conceived of as being non-symbolical.

For instance, AI-researcher and roboticist Rodney Brooks goes as far as to deny internal symbols any existence in his account of embodied cognition, and he claims that this insight is a prerequisite for building human-like intelligence successfully into robotic devices (Brooks, 1999).

In a similar vein, Tangible and Embodied Interaction is promoted as a subfield of interaction design that sees non-symbolic spatial and physical structures derived from sensori-motor experience as predetermining how people understand and make sense of the material world (Hurtienne & Israel, 2007). Under this view, Lund and Waterworth (1998), for instance, define embodied interaction as aiming towards building spatial structures into interfaces rather than seeking to communicate mental models (as suggested by Norman, 1988).

However, some proponents of embodied interaction are not convinced that ascribing primacy to spatial structures over symbolic representations is a tenable strategy. Dourish defines embodied interaction as a design research program that deals with the *relationship* between bodily and physical interaction on the one hand, and a symbolic realm on the other hand. Further, in his insightful field guide to Embodied Cognition (EC), Anderson (2003, p. 100) posits that:

> it is a vice too often indulged by scientists working in EC to make the absence of representations a touchstone of virtue in design, and to therefore suppose, just as do the creatures they devise in the lab, so too must humans display an intelligence without representation. Representation-phobia is a distracting and ultimately inessential rhetorical flourish plastered over a deep and powerful argument. For rather than targeting representation *per se*, the central argument of EC instead strikes at their nature and foundation.

Anderson argues convincingly that the conclusion to be arrived at is not that the symbolic representations of traditional HCI ought to be given up altogether, but rather that we must find ways to systematically relate the symbols and rules of abstract reasoning to the more evolutionary primitive mechanisms which control perception and action. If it can be shown that symbolic representations are at play in the cognitive processing involved in bodily and perceptual experience, then the representation-phobia of embodied interaction must be abandoned.

To expel representation-phobia from situated cognition requires a slightly different therapy. More specifically, it must be shown that representations play a genuine role in online cognitive activity and for the realization of actions. Interestingly, it is worth noticing that in the recently republished edition of *Plans and Situated Actions*, Suchman (2007) actually asserts that viewing her situated cognition framework as non-representational would be a gross misinterpretation of her original argument. In her book, she did not wish to deny the existence of plans conceived of as mentally projected representations of courses of future action. Rather, in 1987, when the book was originally published, Suchman wanted to point towards the critical and, at that time, largely overlooked role played by the immediate context as actions are realized, that is, how real-time requirements and unexpected changes in the context usually call for a continuous revision and modification of the mentally projected plans. Thus, Suchman sums up that her emphasize "is *both* on the utility of projecting future actions *and* the reliance of those projections on a further horizon of activity that they do not exhaustively specify" (Suchman, 2007, p. 19). And further on:

> My position then and now has been that plans are conceptual and rhetorical devices […] that are deeply consequential for the lived activities of those of us who organize our actions in their terms. Just how plans are consequential for the actions they project defined, at least potentially, a territory of mutual interest for the social and cognitive sciences. (Suchman, 2007, p. 20)

Try to compare Suchman's theoretical afterthoughts with the following description taken from a special issue of the journal *Cognitive Science*, where Vera and Simon (1993, p. 10) defend the symbol system approach against the claim that it is incapable of relating inner symbols to actions in everyday settings:

> Symbol systems can be (and sometimes are) used to store in memory representations of external stimuli. They can manipulate these representations as one way of planning actions, and can then execute these actions to change the external situation. Of course, the internal representation of a real scene will be highly incomplete and may be inaccurate, with the result that the actions may or may not have their desired consequences.

Ignoring the fact that the term "system" would undoubtedly be foreign to Suchman's vocabulary, we might say that for Suchman as well as for Vera and Simon, the projection of plans and the execution of actions constitute an open-ended feedback loop between a context, a conceptual system and a sensori-motor system. Hence, the situated action approach should not, as Suchman is careful to underline, be seen as antithetical to the symbolic approach. On the contrary, the two approaches would ideally be able to complement each other in productive ways.

Admittedly, as almost everybody else, Suchman has reservations concerning the narrow algorithmic definition of plans defended by Simon (Newell & Simon, 1972; Simon, 1969). But if the notions of plans and mental representation, as found in the symbolic approach, are broadened so as to include a more dynamic, flexible and situated understanding of the interplay between mental representations and real-time interaction, then a representational view of online cognition might even enrich the SA account of this interplay. This however would require a re-conceptualization of fundamental ideas and concepts belonging to the symbolic approach. The question is of course whether such a re-conceptualization exists? If it can be shown that it does, then the representation-phobia haunting much current work in situated action would seem unjustified.

2. Levels of mental representation

In this section I aim to show that mental representations permeate our perceptual and bodily interaction with physical objects. This presupposes of course that one has a more elaborate notion of what a representation is.

Taken in its most basic sense, a representation involves a sign that stands for something else in a real or imaginary world to the individual. Note that a representation thus consists of a relationship between three elements: world, sign and an agent. Smoke rising in the air stands for fire, but not in itself. It relies on the interpretive work of the perceiver. Just as traces in the forest standing for the feet of an animal or lines and figures on a picture plane standing for the naked body of the model.

Theories of representation in design research have often been criticized for naively believing that representations acquire their meaning from the things or events in reality that they refer to (cf. Krippendorff, 1992). But this criticism seems to have confused “reference” with “representation”. In semiotics, the term ‘reference’ is generally used to designate the relationship between a sign and its referent, while representation is seen as the activity of knowledge-making which the use of signs allows humans to carry out (Nöth, 2000, pp. 148-9). Representations are not mere copies or faded reflections of reality; they involve an active and deliberate use of signs motivated by the individual’s egocentric purposes and cognitive abilities. This is a key insight in the semiotics of Charles S. Peirce (Peirce, 1932). What makes Peirce’s semiotic project seem highly relevant even today is that he anticipates the idea in cognitive science that representation is the most central explanatory device for human cognition (Apel, 1981; Innis, 1994).

When applying the notion of representation to human cognition, we are postulating that mental states somehow act as representations. In cognitive science, it is common, as Cummins (1989) has shown, to distinguish between four different forms of mental representation. Mental representation may thus refer to (1) neuro-physiological states, (2) perceptual content, (3) mental concepts; and (4) more complex symbolic structures, which can take the form of mental models, frames or scripts.

The revolt against representations in embodied interaction hinges upon the idea that human thinking emerges primarily from sensori-motor interaction between an organism and its environment and that representation play no role at all at this level (Johnson & Lakoff, 2002, pp. 249-50). But what if neuro-physiological states in the sensori-motor brain actually show signs of representation themselves? Indeed, the recent discovery of canonical neurons asks us to consider this possibility.

2.1 Sensori-Motor Representations

Canonical neurons are neurons in the pre-motor area of the monkey brain (F5) that are activated when monkeys engage in action execution. In an experiment conducted by Murata et al. (1997), it was shown that canonical neurons that become activated when monkeys physically grasp an object also become active upon the mere observation of the object. Later studies have indicated that similar mechanisms exist in the human brain. This has been taken as evidence for the idea that canonical neurons are crucial for our ability to recognize object affordances: "when a three-dimensional object is seen, the F5 motor neurons for the actions it affords are activated" (Sahin & Erdogan, 2009).

What is interesting about this study is that it points toward the centrality of representations in the sensori-motor system. If canonical neurons fire upon the mere sight of an object, then it means that, on a neuro-physiological level, they inform the organism about a possible future action that can be performed with this object. They do not refer to the object as such, but help to select one functional value of the object that might be relevant for the organism-environment interaction. In this sense, canonical neurons seem to be a good candidate for a neuro-physiological state that acts as a representation as shown in Fig. 1:

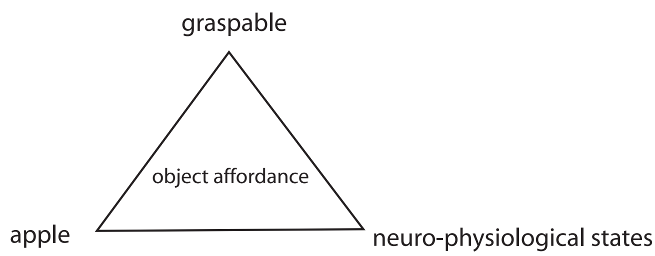

Fig. 1 The representational structure of object affordances

This is in line with Jeannerod (1994) who argues that these neurons are rough motor representations encoded into our semantic knowledge about the object. These motor representations contain initial plans for generating low-level kinesthetic operations presupposed by the execution of an action, for instance grasping, holding, throwing, and so forth. It is important to note that the motor neurons represent a *type* of action and not the bodily operations themselves. This means that if we see a red apple, we may recognize immediately – thanks to the neurons – that it is graspable. But these motor representations evoked from this perception say nothing about whether we should use our left or right hand for carrying out this action.

2.2. Affordances as representation

The discovery of canonical neurons seems to run counter against the concept of affordance, as it was originally introduced in Gibson's ecological theory of perception

(Gibson, 1977). According to Gibson, affordances are possibilities for action that we simply pick up at a pre-reflective level of awareness. As such affordances are thought to be perceived directly without the interference of mental representations or higher-order cognitive operations of any kind. However, Rowlands (2006) has argued that we need to take a refreshing new look at this assumption, because it rests upon a too narrow view of the forms representations may take.

According to Rowlands, there are three kinds of actions that we need to be able to distinguish from each other: *actions, deeds and doings*. An action is inextricably bound up with our intentions such as our volitions, motives or belief-desire couplings. Deeds are defined as pre-intentional acts that depend on the online, feedback-modulated adjustments we perform during the course of interacting with the environment; and doings are sub-intentional acts that are not performed for any reason. Let me try to explain this in more detail.

Actions are performed for some reason in order to accomplish a goal or realize a state of affair that will satisfy our intentions. In philosophy, there is a strict concept of action that holds that the status and identity of an action can only be defined in terms of an antecedent intention (Rowlands, 2006, pp. 96-7). For instance, I may whistle a melody while, at the same time, tap my foot on the floor. Are there one or two actions involved here? According to the strict concept of action this can be determined by asking what intentional states I am the subject of. If I am the subject of a single intentional state then there is only one action. Thus, if my intention is to tap my foot while whistling (for instance as a way to keep the beat), then the corresponding action is one rather than two. If, on the other hand, I am trying to tap my foot and trying to whistle, which just happens to be synchronous, then according to this view I would be performing two actions instead of one (cf. Rowlands, 2006, p. 96).

By reaching out for a goal or state of affair in the near future an intention clearly bears the mark of a representation as it has traditionally been defined. More specifically, an intention can be seen as a plan that is projected imaginatively (however short-lived) and that contains motor representations for how this plan could be fulfilled .

It is important not to confuse these motor representations with the motor representations that have been discovered to be at work at the level of canonical neurons. Compared to intentional states, canonical neurons are intimately bound up with their connection to the physical environment and the perceptual acts we perform in order to cope with our surroundings. Their firing depends utterly on the directness and presence of objects and information available from the environment. Suppose my sister gets angry with me and wants to throw an object at me. She may quickly scan the room visually for objects with the functional value of being *throwable*. In this scanning process canonical neurons may fire upon the sight of an apple, a tennis ball, or a pencil, but not on her viewing the teak dining table in the corner or the car parked outside the window. This means that the canonical neurons are activated because of an action potential evoked from my sisters perceiving of the first three items. Since throwable is indicated from her perception of very different phenomena, it cannot be reduced to some physical properties. Rather it is dependent on the relationship between the surface of the object and the visuo-motor operations of the perceiver. In this sense throwable counts as an affordance. However, since throwable can also be applied to various different phenomena, it must have a certain type-like quality about it. Hence, it also qualifies as a representation.

According to Rowlands many of the acts we perform in order to exploit and manipulate object affordances are not anticipated by any intentional states in the strict sense. Suppose I have noticed the intention of my sister and I therefore start running to escape her attack. In a second or so, she must pick up a throwable object, before I run out of the door. She immediately grasps the pencil and throws it at me. Less than a second is not sufficient to form the intention of whether she should pick up the pencil by grapping it with her whole hand or if grasping it between her thumb and index finger gives a better hold. “In such situations we simply find ourselves acting” (Rowlands, 2006, p. 104). Since these acts cannot be determined with reference to an antecedent intentional state, they are not actions taken in the strict sense. Rather they are what Rowlands calls “deeds”. Deeds are pre-intentional, meaning that even though they are performed in order to fulfill a goal, they cannot be accounted for in terms of intentions. However, they are not reducible either to mere bodily movements, since we do perform them for a reason.

This is what distinguishes deeds from doings. Doings are sub-intentional acts we perform without any reason. If you start attending to your bodily limbs and organs you’ll notice that almost every part of you are in motion: your tongue, toes, eyes, and so forth. We simply cannot help performing these micro-sensational acts or doings to use the term suggested by Rowlands.

By using Rowlands’ three categories, it has been possible to differentiate between representations at the level of actions and representations at the level of deeds, that is the online feedback-modulated adjustments we perform in order to handle everyday situations. These representations are to a very large extent detected by canonical neurons in the flow of our interaction with the physical environment. However, while canonical neurons may inform us about the type-like actions potentials of visible objects – for instance that an apple is graspable – another question that seems relevant to ask is whether we also know from the canonical neurons that the apple is tasty and juicy inside? (cf. Eco, 2000) This seems highly unlikely. The background knowledge we use for inferring such gustative qualities from the apple depends rather on knowledge acquired through our previous experiences of objects of the same kind. If you have never eaten an apple before, you wouldn’t now what to expect. This implies that we draw on prototypical conceptual representations of objects stored in our memory in order to project plans for more complex embodied interactions with the objects at hand. These representations are involved in the forming of intentional states and in the next sections I will briefly account for these representations in terms of concepts, perceptual judgments and mental frames.

2.3 Mental Concepts

It is generally assumed in cognitive science that we use mental constructs of the kind that is usually referred to as ‘mental concepts’. Thus we have concepts for almost everything in this world: apples, stones, the sun, God, shopping malls, and so on. When we encounter things and events, we not only use mental concepts for determining the nature of objects. We also employ them for making what is known as *perceptual judgments and for performing more abstract forms of reasoning*. According to Eco (2000, p. 63) a perceptual judgment is a hypothetical inference based on object affordances and other perceptual inputs that we set up in working memory for purposes of local understanding. Some examples of perceptual judgments would be: “This X is an apple” or “Apples are juicy inside”. Or “This X in front of me is a stone”, “It is struck by sun light”, and “it is hot” . We may even construct more complex propositional representations out of these perceptual judgments, for instance, if we manage to reach the conclusion that “The sun heats the stone”. In this instance we subsume a particular relationship between two entities in the world under the general law of C*ause-Effect*.

Interestingly, Johnson (1987, pp. 37-40) has argued that such logical inference patterns like Cause-Effect, the Law of the Excluded Middle and the Law of

Transitivity, which we normally see as part of abstract symbolic reasoning, have their intuitive basis in our daily experiences with physical objects and orientation in space. The idea then is that symbolic reasoning is constrained and, to some extent, made up from structures derived from sensori-motor interaction between the organism and its environment (this idea is accurately captured in the so-called Spatialization of Form Hypothesis, see Lakoff 1987, p. 283). While Johnson who is an esteemed philosopher argues convincingly for this idea, the truth of his claims has in fact not only been anticipated, but also demonstrated by the experimental phenomenologist Albert Michotte. Through a series of empirical experiments Michotte found that our immediate sensory experience consists not only of objects or entities, but also their causal relations (see e.g. Michotte, 1963). For example, in one of his studies Michotte proved that the duration of contacts between two colliding objects is crucial for how we understand their causal relation. This study is nicely summarized by Heft (2003, p. 167):

> If an object "makes contact" with another, and then ceases moving just as the second object begins to move along the same prior trajectory, observers report that the first object "launched" the second. However, when the first object even after contact continues to move along the trajectory with the now-moving second object, observers report a pushing or a chasing ("entraining") rather than a launching effect.

Michotte's study shows that spatio-temporal relations among objects that we experience might lead to our perception of different types of causal effects. A result like this challenges the influencing assumption that was introduced by Hume, namely that entities in sensory experience are disjoint and that the order and lawfulness, they appear to have, are the result of reflective operations on the part of the subject that subsume sensory tokens under *a priori* logical principles of the intellect (an idea that reached its culmination in Kant's *Critique of Pure Reason*). What Michotte's experiments reveal is that some causal effects are not the result of reflective intellectual judgments, but of pre-reflective perceptual judgments.

However, in order to provide an adequate account of human cognition, we need to move beyond the level of perceptual judgment toward inferences and mental representations of an increasing complexity. There is clearly a difference between perceiving one ball as having a launching effect on a second ball and the causal relations that may be involved, for instance, in the planning of a trip to Montreal. Eco (2000) has suggested that we conceived of inferences as semiotic processes that work their way up from the level of sensory experience to higher-order mental representations as a kind of scaffolding process. Like a cognitive spiral where low-level meaning structures develop into ever more complex mental representations. For instance, according to this model object affordances are integrated as part of the perceptual image of 3-dimensional objects; this perceptual content can then integrated as part of perceptual judgments or be subsumed under mental concepts, when, for instance, we identify this X as a species belonging to this category. Further, mental concepts can be integrated as part of more advanced symbolic forms such as mental models, frames and scripts. If we take an apple as our example this scaffolding process can be diagramed as in Fig. 2:

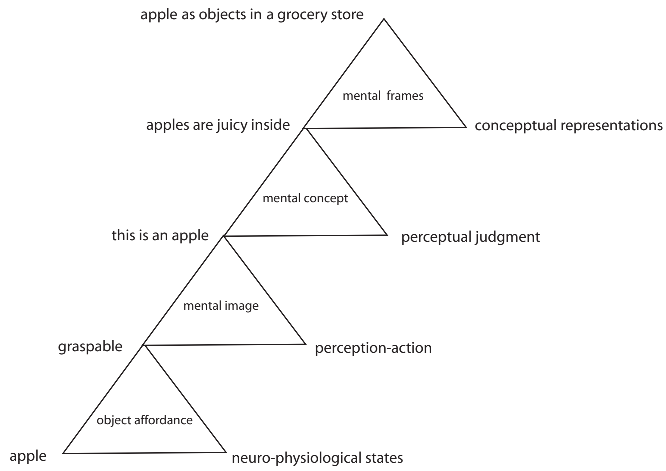

Fig. 2 Diagram of how higher-order cognition relates to the mechanism controlling perception and action.

The interconnected triangular structure of Fig. 2 illustrates how representations from low-level cognition grow into more complex knowledge representations. The next two sections will explain the nature of these representations.

2.4 Mental Frames

Apart from perceptual judgments, mental concepts are also important for building up even more complex networks of conceptual representations. Among such networks we find mental frames. Basically, a frame is conceived as referring to an organized system of concepts representing stereotypical background knowledge that people unconsciously draw upon when using language and thinking and acting in the world (Kövecses, 2006, p. 69). Consider Fillmore's (1982) so-called BUY frame (see Fig. 3). The idea is that we would not be able to understand the meaning of the verb "buy", unless we automatically activate a mental frame that allows us to make default assumptions about the situation that the verb can be used to describe. For example, "buy" presuppose a *buyer* owing an amount of *money*, a *seller* that offers some *goods* that the buyer wants to have, and so forth (cf. Ungerer & Schmid, 2006). Since the frame thus represents intentional roles they are important for forming goal-oriented plans.

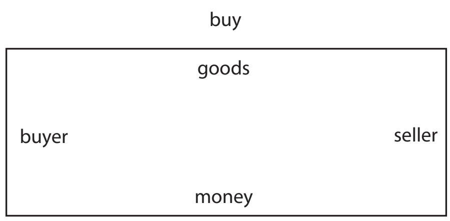

Fig. 3 Fillmore's Buy Frame (adapted from Fillmore, 1977)

In Artificial Intelligence Minsky (1975) found the notion of frame valuable for explaining people's default knowledge about objects or events involved in routine tasks and reoccurring situations in work places. In cognitive psychology and cognitive linguistics it has also been shown that we use frames in our perception, planning and memory for events (Coulson, 2001; see also Barsalou, 1992). In so doing, we become able to create expectations and make predictions about the consequences of actions in a given context.

For example, I may have some friends coming over for dinner. To surprise them I want to make a juicy apple pie. But I realize that I don't have any apples at home, so I go to the grocery store to get some. To fulfil this intention I must draw upon a rich array of conceptual representations and knowledge structures. For instance, I need to be able to distinguish certain sort of apples from others, so that I'll get the right ones for my pie. At the point where I have to pay for the apples, it might be assumed that I draw unconsciously on a BUY frame in order act appropriately in this economic transaction. Obviously, the structure and content of such a frame will be sensitive to people's socio-cultural learning and background. For instance, in the Southern parts of Europe it is custom to bargain over the prize, while in the Scandinavian countries it is not. Hence, the precise nature of the BUY frame must be assumed to vary slightly from one cultural context to another.

However, while this is relatively uncontroversial, one of the main critiques of the notion of frames is that it is unable to account for how people manage to cope with changes in practices, unforeseen occurrences or novel situations that violate their expectations. Being products of long-term memory, frames offer pre-existing and default mental templates for thinking, not dynamic and flexible models adjustable to the immediate context as actions are realized. This is one of the central objections made by proponents of situated action and distributed cognition.

To many situated activity theorists, our everyday life seems most of all to resemble river rafting. In river rafting it is impossible for the canoeist to act according to abstract mental plans, because the violent stream, turbulence and protrusions, force him constantly to improvise and act according to *ad hoc* rules in order to control and keep the canoe in balance.

Everyday situations have their own turbulences and protrusions. Ethnographic field studies of people's shopping behaviour have been taken as evidence for this. When people are going to put the items they have bought into a shopping bag, there are typically a whole series of factors that might distract them from the purpose of their actions. Items come down the line in the shopping mall in a random order, so it is a challenge to us to decide what items should go in the bottom and which ones at the top. Then, while we are in the midst of packing our bag, our mobile phone may start ringing, or the cashier may interrupt the course of actions telling us that we haven't given him the exact amount of money. In such instances, it is unlikely that we rely on an abstract representation or algorithm in the head telling us what to do. Or put more accurately: perhaps there is such an algorithm but it is of no use, since unforeseen occurrences and events constantly force us to deviate from such idealized mental models. To explain how we nevertheless succeed most often in getting home without finding splashed eggs in our shopping bag, it has been suggested that we use the items themselves as external vehicles for informing us about weight and size and whether they should be placed in the bottom and at the top. By externalizing cognitive resources into the environment we are able to adjust to real-time requirements (see e.g. Hutchins, 1995; Kirsh, 1995).

However, even though I agree that such situations undoubted question the universal validity of abstract mental representations such as frames, they should not be taken as a model for all forms of cognition (Rowlands 2006). Just as well-structured intentional acts such as following a recipe or going to the grocery store should not be taken as a model either. Moreover, even stressful situations do not prove that the role traditionally assigned to representations can be completely taken over by actions without representations. Spontaneous acts and improvisation that we may perform during our shopping do not unfold beyond the representational. If so, they would be reducible to mere doings, bodily movements that are performed without any reason.

This is clearly not the case. We perform those acts spontaneously in order to accomplish goals or satisfy our will. Hence, they are deeds that – as we have seen – rely to a large extent on representations, however not of the kind that intentions are made up from. Deeds stand halfway between doings and actions. They involve our real-time coping with representations at the level of canonical neurons and direct perception.

3. Conclusion

To sum up, in this paper I have argued that representations permeate our entire existence from its most basic manifestation at a neuro-physiological level to conceptual representations in people's memory motivated by social-cultural learning and practices. We must of course recognize that representations come in various forms defined by increasing complexity (cf. Fig. 2).

If this idea is taken seriously, then we must reject the idea that the absence of representation should be made a touchstone of virtue in design. Many representation-phobic researchers either from the design or HCI community seem to believe that embodied interaction is a design strategy for working primarily with physical structures at the sensori-motor level. However, from a representational point of view, physical structures can only be properly understood if they are seen as one element in a triadic relationship between object, sign, and organism.

Taking my point of departure from Anderson (2003) and Suchman (2007) I initially claimed that the central argument of embodied cognition and situated action consists in a re-conceptualization rather than a rejection of representation. More specifically, this re-conceptualization relies on one being able systematically to relate the symbols and rules of abstract reasoning to more evolutionary primitive mechanisms, which control perception and action (Anderson, 2003). In order to carry out this task I first drew upon Rowlands' (2006) distinction between actions, deeds and doings, which allowed me to isolate representations at the level of intentions from representations at the level of our pre-intentional perceptual interaction with the physical world, that is deeds. The latter are adjustable to dynamic changes in the context. In this sense, it seems as if deeds could have the potential of bridging a situated action perspective with a representational perspective. This however needs to be studied more carefully in a future work.

Furthermore, I have attempted to show how the ability to infer generality in terms of laws and relations between objects in perception (such as Cause-Effect), which we normally associate with abstract reasoning, in fact emerge from primitive mechanisms in perception itself.

In order to provide the reader with a clearer view of the rich interplay that are at stake from neuro-physiological states to abstract thinking, I have laid out a diagram in Fig. 2. Much of what has been said about the upper levels in this diagram is standard knowledge in cognitive science. However, the idea of including these levels in the diagram is of course that embodied interaction encompasses the whole continuum of representations from canonical neurons to mental frames. It is this continuum that accounts for the full complexity of embodied interaction.

References

Anderson, M. L. (2003). Embodied cognition: A field guide. *Artificial intelligence*, *149*(1), 91-130.

Apel, K. O. (1981). *Charles S. Peirce: From pragmatism to pragmaticism*. Univ of Massachusetts Pr.

Barsalou, L. W. (1992). Frames, concepts, and conceptual fields. *Frames, fields, and contrasts: New essays in semantic and lexical organization*, 21-74.

Brooks, R. A. (1999). *Cambrian intelligence: The early history of the new AI*. The MIT Press.

Bærentsen, K. B., & Trettvik, J. (2002). An activity theory approach to affordance. In *Proceedings of the second Nordic conference on Human-computer interaction* (p. 60). ACM.

Carroll, J. M. (1987). *Interfacing thought*. MIT Press Cambridge, MA.

Carroll, J. M. (1991). *Designing interaction: Psychology at the human-computer interface*. Cambridge Univ Pr.

Coulson, S. (2001). *Semantic leaps: frame-shifting and conceptual blending in meaning construction*. Cambridge Univ Pr.

Dourish, P. (2004). *Where the action is: the foundations of embodied interaction*. The MIT Press.

Eco, U. (2000). *Kant and the platypus: essays on language and cognition*. Mariner Books.

Ehn, P., & Linde, P. (2004). Embodied Interaction–Designing beyond the physical-digital divide. *Malmö, School of Arts and communication, Malmö University*.

Fillmore, C. J. (1982). Frame semantics. *Cognitive linguistics: basic readings*, 373–400.

Gibson, J. J. (1977). The theory of affordances. *Perceiving, acting, and knowing: Toward an ecological psychology*, 67-82.

Heft, H. (2003). Affordances, dynamic experience, and the challenge of reification. *Ecological Psychology*, *15*(2), 149-180.

Hornecker, E. (2005). A design theme for tangible interaction: embodied facilitation. In *ECSCW 2005* (pp. 23-43). Springer.

Hurtienne, J., & Israel, J. H. (2007). Image schemas and their metaphorical extensions: intuitive patterns for tangible interaction. In *Proceedings of the 1st international conference on Tangible and embedded interaction* (p. 134). ACM.

Hutchins, E. (1995). How a cockpit remembers its speeds. *Cognitive Science: A Multidisciplinary Journal*, *19*(3), 265-288.

Hutchins, E. (2005). Material anchors for conceptual blends. *Journal of Pragmatics*, *37*(10), 1555-1577.

Innis, R. E. (1994). *Consciousness and the Play of Signs*. Indiana Univ Pr.

Jeannerod, M. (1994). The representing brain: Neural correlates of motor imagery and intention. *Behavioral and Brain Sciences*, *17*, 187-245.

Johnson, M. (1987). *The body in the mind: The bodily basis of meaning, imagination, and reason*. University of Chicago Press Chicago.

Johnson, M., & Lakoff, G. (2002). Why cognitive linguistics requires embodied realism. *Cognitive Linguistics*, *13*(3), 245-263.

Kaptelinin, V., & Nardi, B. A. (2006). *Acting with technology: Activity theory and interaction design*. MIT Press Cambridge, MA.

Kirsh, D. (1995). The intelligent use of space. *Artificial intelligence*, *73*(1-2), 31-68.

Klemmer, S. R., Hartmann, B., & Takayama, L. (2006). How bodies matter: five themes for interaction design. In *Proceedings of the 6th conference on Designing Interactive systems* (p. 149). ACM.

Krippendorff, K. (1992). Transcending semiotics. *Objects and images*, 24–48.

Kövecses, Z. (2006). *Language, mind, and culture: a practical introduction*. Oxford

University Press, USA.
Lakoff, G. (1987). *Women, fire, and dangerous things: What categories reveal about the mind*. University of Chicago press Chicago.
Lakoff, G., & Johnson, M. (1999). *Philosophy in the flesh: The embodied mind and its challenge to Western thought*. Basic Books.
Lund, A. (2003). *Massification of the intangible: An investigation into embodied meaning and information visualization*.
Lund, A., & Waterworth, J. A. (1998). Experiential Design: Reflecting embodiment at the interface. In *Computation for Metaphors, Analogy and Agents: An International Workshop, University of Aizu, Japan*.
Michotte, A. (1963). *The perception of causality*. (T. R. Miles, Tran.). Methuen London.
Minsky, M. (1975). Frame-system theory. *Thinking: Readings in cognitive science*, 355-376.
Murata, A., Fadiga, L., Fogassi, L., Gallese, V., Raos, V., & Rizzolatti, G. (1997). Object representation in the ventral premotor cortex (area F5) of the monkey. *Journal of Neurophysiology*, *78*(4), 2226.
Nardi, B. A. (1996). *Context and consciousness: activity theory and human-computer interaction*. The MIT Press.
Newell, A. (1980). Physical symbol systems. *Cognitive Science: A Multidisciplinary Journal*, *4*(2), 135-183.
Newell, A., & Simon, H. A. (1972). *Human problem solving*. Prentice-Hall Englewood Cliffs, NJ.
Norman, D. A. (1988). *The psychology of everyday things*. Basic books.
Nöth, W. (2000). Handbuch der Semiotik [2, vollständig neu bearbeitete und erweiterte Auflage]. *Stuttgart: Verlag JB Metzler*.
Payne, S. J. (2003). Users' mental models: The very ideas. *HCI Models, theories and frameworks: Toward a multidisciplinary science*, 135–156.
Peirce, C. S. (1932). *Collected papers of charles sanders peirce*. (C. Hartshorne & P. Weiss, Eds.). Harvard University Press.
Rowlands, M. (2006). *Body language: representation in action*. The MIT Press.
Rowlands, M. (2009). Situated Representations. In P. Robbins & M. Aydede (Eds.), *The Cambridge Handbook of Situtated Cognition* (pp. 117-133). Cambridge: Cambridge University Press.
Sahin, E., & Erdogan, S. T. (2009). Towards linking affordances with mirror/canonical neurons. ISCIS Special Session on Cognitive Cybernetics and Brain Modeling. Retrieved from http://www.rossiproject.net/publication.html
Schank, R. C., & Abelson, R. P. (1977). *Scripts, plans, goals and understanding: An inquiry into human knowledge structures*. Lawrence Erlbaum Associates Hillsdale, NJ.
Simon, H. A. (1969). *The sciences of the artificial*. The MIT Press.
Sinha, C. (2005). Blending out of the background: Play, props and staging in the material world. *Journal of pragmatics*, *37*(10), 1537-1554.
Suchman, L. A. (1987). *Plans and situated actions: The problem of human-machine communication*. Cambridge Univ Pr.
Suchman, L. A. (2007). *Human-machine reconfigurations: Plans and situated actions*. Cambridge Univ Pr.
Varela, F. J., Thompson, E., & Rosch, E. (1991). *The embodied mind*. MIT press Cambridge, MA.
Vera, A. H., & Simon, H. A. (1993). Situated action: A symbolic interpretation. *Cognitive Science: A Multidisciplinary Journal*, *17*(1), 7-48.
Wilson, R. A., & Clark, A. (2005). How to situate cognition: Letting nature take its course. *The Cambridge Handbook of Situated Cognition*.
Winograd, T., & Flores, F. (1986). *Understanding computers and cognition: A new foundation for design.* Ablex Publishing Corporation.

Zlatev, J. (2005). What's in a schema? Bodily mimesis and the grounding of language. *From perception to meaning: Image schemas in cognitive linguistics*, 313.