

“Embodying” the Internet: Towards the Moral Self via Communication Robots?

Johanna Seibt · Marco Nørskov

Received: 31 July 2011 / Accepted: 10 January 2012 / Published online: 19 February 2012
© Springer-Verlag 2012

Abstract Internet communication technology has been said to affect our sense of self by altering the way we construct “personal identity,” understood as identificatory valuative narratives about the self; in addition, some authors have warned that internet communication creates special conditions for moral agency that might gradually change our moral intuitions. Both of these effects are attributed to the fact that internet communication is “disembodied.” Our aim in this paper is to establish a link between this complex of claims and past and ongoing research in phenomenology, empirical psychology and cognitive science, in order to formulate an empirical hypothesis that can assist development and evaluation of recent technology for *embodied* telecommunication. We first suggest that for the purposes of interdisciplinary exchange, personal identity is formally best represented by a selection function that (for temporal intervals of variable length) “bundles” capacity ascriptions into identificatory narratives. Based on this model, we discuss which cultural changes engendered by the internet affect the construction of personal identity in ways that diminish our ethical sensitivities. In a second step, working from phenomenological claims by Martin Buber, we argue that disembodied communication severs two modes of cognitive function, preconceptual and conceptual, which tie together moral motivation, self-experience, and identity construction. We translate Buber’s claims into the theoretical idiom of the “theory of cognitive orientation,” a psychological theory of motivation that links up with recent research in embodied cognition. In a third step, we investigate whether the embodiment of the internet with communication robots (e.g., telenoids) holds out the prospect of reverting this structural change at least partially. We conclude by formulating an empirical hypothesis (for researchers in cognitive science) that has direct import, we submit, on the question whether embodied telecommunication promises a new form of ethically sensitive self-constituting encounter.

J. Seibt (✉) · M. Nørskov
Department for Philosophy and the History of Ideas, Aarhus University, Aarhus, Denmark
e-mail: filseibt@hum.au.dk

M. Nørskov
e-mail: filmnp@hum.au.dk

Keywords Personal identity · Embodied internet communication · Teleoperated communication robots · Cognitive orientation · Embodied cognition · Philosophy of dialogue · Second person cognitive science · Moral cognition

Internet communication technology has changed our communicational practices in ways that seem to affect our self-understanding. Are we losing our sense of self in the multiplication of our “identities” or self-presentations on the internet, and if so, are we losing the very ground of moral agency? In view of the multiplication of self-presentations, some authors have called for a critical reflection of the ethical and political implications of a transition from the traditional Western understanding of the self as a “unified unchanging core” into the direction of non-Western relational notions of self (e.g., Ess 2009, 2010). Others, however, hold that the multiplication of “identities” per se amounts to a problematic assault on moral subjecthood, as “loss of authentic individuality” that leads to a “deindividuation or reduced sense of responsibility” (Csepeli 2009: 112). For some, the ethically problematic aspects of the internet pertain to the fact that internet “identities” or self-presentations are a product of controlled fabrication rather than exposed involuntarily (e.g., Cocking 2010). Others emphasize that it is the anonymity of the internet that fundamentally alters the conditions for moral agency, to the extent that our moral intuitions in real life may become affected (Csepeli 2009; Dreyfus 2001: 73ff, Turkle 2011: 211ff). Frequently, the ethically problematic aspects of internet communication are simply attributed to the “lack of physical contact,” i.e., to the fact that internet communication is “disembodied.” The details of this causal attribution, however, are either taken to be simply “understood” or remain at the level of providing pointers to research on “body language” and the communicative role of facial expressions.

What is it about “disembodied” internet communication that could disturb moral cognition and the formation of our moral self-understanding? Can these effects be mitigated? In view of recent attempts at adding elements of embodied communication into internet communication, e.g., by teleoperated communication robots that can “hug”, it seems important, in our view, to determine more precisely in which sense aspects of physical contact in direct personal encounters could be credited with creating the best conditions for the formation of a self-understanding that is open to the ethical demands of the situation. Our aim in this paper is to pave the way for targeted empirical research on this issue. By linking research on the phenomenology of direct personal communication (“dialogue”) with empirical studies on the formation of motivations in cognitive psychology and ongoing research on “embodied cognition” in cognitive science, we formulate an empirical hypothesis that can, as we try to show by an illustration, prove useful for the development and evaluation of recent technology for “partially embodied” telecommunication.

We proceed as follows. Given the fluidity of current terminology, where some authors speak of “multiple identities” and others of “plural selves,” we first suggest that—for the purposes of interdisciplinary exchange—personal identity is best represented by a selection function that generates identificatory narratives. On the basis of this model one can distinguish more clearly between three types of claims about the effects of the internet on our “identity.” For the context of our project, here we set aside claims that pertain to changes in the possible *content* of identificatory narratives

of groups and persons. Of primary concern for our purposes are claims that specifically address the internet as a medium of “disembodied” communication and associate with it certain changes in the *structure* of our personal identificatory narratives. We discuss several candidates for such structural changes and find that none of these *directly* implies a change in the conditions of moral agency. That the relevant changes could foster a degradation of our moral self-understanding would follow only in combination with an account of moral cognition in embodied communication.

To such an account we turn in the second section and ask why embodied communication could be said to hold the key to our moral sense of self. The claim that embodied communication supports ethically sensitive identity constructions can be argued, we suggest, in a two-step procedure. The first—and mainly heuristic—step consists in a phenomenological investigation of direct personal communication. We sketch Martin Buber’s observations that “dialogue” engenders a special mode of cognition that supports ethically sensitive identity constructions or the formation of a moral self. In a second step we reframe these insights in terms of the “theory of cognitive orientation,” an older approach in empirical psychology, in order to establish links to current research to embodied cognition. Based on these links and the phenomenology of moral judgment, we formulate the empirical hypothesis that *direct personal communication foregrounds a pre-conceptual mode of cognition (“intensified orientation”) that furthers ethical sensitivity.*

In the third section we report some first (“anecdotal”) evidence from a pilot study conducted in March 2011 in Denmark that would seem to suggest that teleoperated communication robots also have the capacity to keep us in a prolonged phase of “intensified orientation” where we are open to the ethical demands of the situation.

We conclude that the ethical discussion of embodied telecommunication via, e.g., communication robots, should not center on the issue of whether human-robot interaction ever could *replace* dialogue among humans. Rather, the discussion should focus on the question whether communication robots can *embody internet communication in ways that support the formation of our ethical sensitivities as much as human dialogue.* To be sure, here philosophers need to await the results of cognitive science on, for example, the empirical hypothesis we formulate in this paper, namely, that direct personal communication prolongs intensified orientation, a mode of preconceptual cognition that ties our self-understanding to ethical concerns. Beyond the specifics of the example discussed, however, we also present this investigation in order to make a case for an escape from the alleged dilemma of either accepting new technology or else staying with traditional practices that the latter never can fully replace. In sum, our aim here is to make a case for a more differentiated approach to the ethical discussion of new technology that focuses on the latter’s interaction potential for accentuating and even strengthening those capacities and dimensions of human practice we do not wish to lose.

1 Personal Identity and Cultural Change: Precisely How Does the Internet Change Who We Are to ourselves?

Internet communication has produced significant changes in our cultural practices, from institutional to economic to social interactions. Both in public media and

research literature there is considerable discussion about whether these changes amount to a loss of a fixed cultural, social, ethnic, national, or even personal “identity.”¹ However, there is no established and broadly accepted theoretical model that could ground the use of the idiom of “identity” in this debate, which frequently appears to be used for its rhetorical momentum only. Thus, we set out with some terminological clarifications which, we hope, will not only be useful for our agenda in this paper—viz to investigate how embodying aspects of internet communication affects our personal “identities” and “sense of self”—but also for other projects pertaining to the relationship between personal identity and cultural change due to technology.

For philosophers an account of “personal identity” is an account of the *numerical* and *trans-temporal* identity (i.e., identity at a time and across times) of an entity that has the *sortal* identity of an individual human being. By definition, none of these identities—numerical, trans-temporal, or sortal (for natural kinds)—can vary over time.² This makes for poor contact with the use of the notion of “identity” in sociology and anthropology, where since (Tajfel 1974; Tajfel and Turner 1979) it has become customary to speak of the “identities” of persons as “changing” in tandem with social or cultural change, of “identity dynamics” and “identity loss.” From the point of view of the philosophical tradition such talk about changing social, ethnic, or cultural “identities” pertains to aspects of the *qualitative* identity of a person, i.e., of who we are in the sense of “what we are like,” our “empirical features”, and such descriptions of the “empirical Ego” (Kant) are taken to be a different issue altogether. More recently, however, philosophers questioned this traditional separation and proposed—following Hume—to define the numerical and transtemporal identity of persons in terms of their qualitative identity or empirical features—e.g., in terms of continuity relations among experiential contents (Parfit 1984). Whether this strategy is ultimately successful we wish to leave open here, but the approach certainly has been helpful to draw attention to the fact that at least our moral self-understanding hinges on our qualitative identity—who we take ourselves to be for the purposes of moral agency depends on “what we are like,” as perceived by ourselves and to others. The question whether and how cultural changes, such as internet communication, “affects personal identity” is frequently understood to pertain to effects on our self-understanding as a unified and permanent self *for the purposes of moral agency*. Thus, in this context at least, we believe that philosophers are well advised to follow the “neo-empiricist line” and to develop an account of personal identity understood as an account of a person’s self-understanding based on the person’s “qualitative identity.” Once we have a model of the formation of self-conceptions, i.e., a conception of who we are to ourselves, we can further determine whether the model is empirically plausible, which type of cultural changes are likely to influence the formation of self-conceptions more deeply than others, and whether it can be argued that some of these latter more profound changes as such have implications for moral agency. In the following paragraphs, we will first sketch such

¹ See e.g., Jensen (2003).

² That these identities are invariant is taken as an ‘axiom’ of ontological research.

a model of personal identity and then show how it can be used to evaluate claims about the consequences of internet communication on our self-conceptions.³

Consider common sense judgments about sameness and difference of persons at and across time, i.e., judgments such as “I am not the person I used to be”, or “by moving to the USA you have become a different person,” or “no matter which relationship Tom’s in, he’s always the same”, etc. These judgments—this we submit here as a phenomenological claim—are grounded in a experience of “what it is like to be X”, where X is the name of a person (“me,” “you”, “Tom”, etc.); this experience is—in its immediate phenomenal presentation—a preconceptual intuition of a complex quality. When we try to make this complex quality explicit and describe who we are to ourselves and who others are to us, we commonly present a collection of valuative commitments. These valuative commitments are closely related to the experiences and capacities of the person in question, as perceived by ourselves or by others.⁴ A properly “data-driven” theory of personal identity must differentiate (a) among a “first-person”, “second-person”, and “third-person” perspective on personal identity, and (b) between experiences and capacities that have been recognized as having identificatory status from a first/second/or third-person perspective and those that have not. When I try to conceptualize and articulate who I understand myself to be, I will mention capacities that may or may not overlap with capacities you mention when you explicate who you understand me to be; and yet another posture of identificatory understanding opens up if we judge personal identity from the distance of a third-person point of view. Briefly, we select experiences and capacities into identificatory sets that form narratives of who we are, to ourselves and to others, and who others are to us. Such narratives consist of elements of the past presented as indications of valuative commitments that will guide future behavior; they may persist through many changes in the experiences and capacities of the person they define, but they also may be modified over time to the extent that the subject of the narrative appears to have become a ‘different person’ to herself/himself or others.

More formally speaking, talk about personal “identity” refers either (a) to a set of experiences and capacities experienced in the mode of action (first-person perspective), or else (b) a set of capacities of another human being experienced in the mode of perception and emotional engagement (second-person perspective), or else (c) to a set of capacities of another human being experienced in the mode of perception (third-person perspective); in each case, these sets form a narrative or collection of valuative commitments that identify a person during a temporal interval. In the

³ For details and applications of this model, see Seibit (2011), where it is used to reconstruct the changes in our self-understanding in value conflicts.

⁴ In the context of this paper, we use the term “capacity” in a wide sense, denoting actual *interaction potentials* of all sorts: natural or acquired, active or passive, cognitive or emotional or practical, intended or unintended, unrealized (such as my capacity of speaking English when I am silent) or realized capacities. In particular, our usage of “capacities” includes ‘practices,’ which are here subsumed as a species of realized capacities. Most importantly, note that on this wide notion of capacity, an experience also counts as a sort of cognitive capacity, next to beliefs and emotions, namely, as a capacity the realization of which is partly active and partly passive. Thus, the conjunction ‘experiences and capacities’ should (here and thereafter) be read as “experiences and *other* capacities,” and it is only for rhetorical reasons that we—in line with the historical debate about personal identity—highlight experiences to emphasize a passive element in the construction of personal identities. For the wider theoretical context of this interactivist view of cognition, see Bickhard (2003, 2010).

following, we sketch the basic idea for the construction of an analytical model of the interaction of sociocultural change and personal identificatory narratives; given the aims of this paper, we focus solely on the first-person perspective, i.e., on the self-ascribed identificatory narratives.

We propose to take personal identity—in the sense of a person’s self-conception—to be a selection function that ranges over capacities (including experiences, e.g., *honest, athletic, grew up in Paris*) and produces personal identificatory narratives as outcomes. We consider this functional representation as the structural model of a cognitive process that has sufficient empirical support to count at least as empirically plausible, as we describe in “Section 2” of this paper. More concretely then, drawing on the notion of capacities set out above, let us note that an individual human being has (a) *sortal capacities* due to being members of a certain kind, and (b) *individual capacities*, among which are her or his (c) experiences. My personal identity is a selection function on my individual capacities (and experiences). Quite analogously, my cultural, social, and ethnic identities are selection functions on my sortal capacities. Unlike my first-person personal identity, however, the selection function for sortal capacities constituting my cultural, social, or ethnic identificatory narratives are determined by a group of individuals, not by myself.

In formal terms, let: $A = \{a_1, \dots, a_n\}$ be a group of individual human beings, T_{a_1} be the set of (possibly overlapping) temporal intervals t_i that together span the past and present lifetime of an individual a_1 in A , let C_{a_1, t_i} be the set of capacities of individual $a_1 \in A$ during $t_i \in T_{a_1}$. Let C_{a_1} be the union of all such temporary capacities, $C_{a_1} = \cup C_{a_1, t_i}$, called the “diachronic capacities” of individual a_1 . Let V_{a_1} be the powerset of C_{a_1} , containing the possible personal narratives or valuative commitments of a_1 . Finally, let C_{ai} be the union of the diachronic capacities of all members of the group A , and V_{ai} the set of possible group narratives for A . The selection functions for a person’s personal identificatory narrative (f_{pi}) and various types of sortal identificatory narrative (s_{si}), i.e., cultural, social, or ethnic “identity”, can then be stated as follows:

$f_{pi}: T_{a_1} \times C_{a_1} \rightarrow V_{a_1}$, with $f_{pi}(t_i, C_{a_1}) = V_{a_1, t_i} \in V$. In other words, personal identity is a function that maps at any given interval the total set of an individual’s past and present capacities to a subset of the latter, forming the individual’s ‘personal narrative’ at that time. A personal narrative is a list of capacities that the individual considers to be representative for who she or he is at that time, and endorses as valuative commitments.

$f_{si}: T_{ai} \times C_{ai} \rightarrow V_{ai}$, with $f_{si}(t_i, C_{ai}) = V_{ai, t_i} \in V_{ai}$. In other words, a person’s sortal identity is a function that maps at any given interval of lifetime of the group the total set of the group’s past and present capacities to a subset of the latter, forming the group’s “sortal narrative” at that time. The group’s sortal narrative (i.e., either its cultural narrative, or its social narrative, or its ethnic narrative) is a list of capacities that the group considers to be representative for who they are, and endorses as valuative commitments.

An individual’s full identity at a time consists of the union of his or her personal and sortal identificatory narratives at that time, and the individual’s (self-ascriptive or “first-person-point-of view”) identity consists in both the individual’s selection functions for personal and for sortal identity.

There are several theoretical advantages of working with selection functions rather than applying the traditional strategies of trying to characterize “identities” in terms of abstract universal features or, strictly extensionally, in terms of total (temporally complete up to the present) sets of capacities. Most importantly, if we model personal identity as selection functions, we can account for stability and change of such narratives as the anthropological and sociological data would suggest: we can allow for the possibility that not every change in my individual experiences and capacities necessarily changes my personal identificatory narrative, and that not every change in sortal capacities necessarily changes my cultural, social, or ethnic identificatory narrative. Furthermore, it is a virtue of our formal representation, we believe, that the formalism does impose any simple dependencies between personal identities and sortal identities—the extent to which changes in sortal capacities propagate into changes in personal identities can thus be stated in as differentiated a fashion as empirical research suggests.⁵

In terms of our model, we can now distinguish three types of theses that are associated with the slogan that the “internet affects our identity.” First, there is the thesis that cultural changes engendered by the internet affect our *sortal identity*, i.e., our identificatory cultural, social, or ethnic narratives.⁶ Second, there is the thesis that the cultural changes engendered by the internet affect our *personal identity*, i.e., the *content* of our self-conceptions. As a general principle, it holds that new elements in the set of sortal capacities C_{ai} of a person (e.g., using a writing system) may alter the set of individual capacities C_{a1} by adding (e.g., making music) or subtracting (e.g., memorizing); the project of the second type of thesis is to show how the agentive possibilities introduced by the internet alter the content of the set of an individual’s capacities $C_{a1, r1}$, engendering some and deleting some.⁷ Third, there is a more specific version of the second claim, namely, the claim that (a) there are cultural changes that are directly due to internet communication *qua disembodied communication* and (b) that these cultural changes effect deep *structural changes* in our personal identities, i.e., (not only the content but also) the *way* in which we form our self-conceptions. In empirical and philosophical studies of the internet, it is not always clear which type of claim is supposed to be supported by the evidence adduced. In the remainder of this section, we will discuss some of the observations that have been, or can be, offered in support of the third claim. For each of these observations, we will also briefly address the question (c) whether and why such modifications of our self-conceptions could be said to have direct implications for our capacity to act morally.

⁵ For example, individual experiences of ethnic conflict may become part of the ethnic identificatory narrative and thus passed on to personal identificatory narratives of the next generation; see Mahmoud (2011).

⁶ For example, social media have been said to blur ethnic identities (Jensen 2003), but create new social identities.

⁷ Turkle (2011) contains an impressive collection of concrete cases documenting how social media have come to change social practices and social etiquette pertaining to the introduction, dissolution, and maintenance of social relationships, and thus also affects our personal narratives on how we perform relative to these norms, whether we take ourselves to be reserved or outgoing, sensitive or pragmatic, compliant or wild, etc.

Let us consider, then, whether and which structural modifications to self-conceptions can be directly attributed to the fact that internet communication is “disembodied”, i.e., that it is largely textual, involves no direct visual or auditory transmissions of face and voice but rather represents the participants semantically by name and self-declarations (avatars, profiles etc.). To refer to this special character of internet communication as disembodied communication we will hereafter speak simply of the “disembodied internet.” That the ‘disembodied internet’ enables us to work with multiple self-presentations in profiles on social media and avatars in game worlds, and even quasi-simultaneously so, is surely the most obvious candidate for an illustration of structural effects on our personal identificatory narratives.⁸ Frequently, these multiple self-presentations diverge from each other merely in the sense of modifications relative to social role and social context—something we are well-familiar with from the times before the internet (Back et al. 2010). But there are sufficiently many instances of socially well-functioning people with successful careers who construct and simultaneously operate with profiles that are widely different from each other and from the personal identity of the user in real life, often including a gender switch. Our new cultural (*qua* technological) capacity of mixing our real and virtual lives “on the go,” continuously and as we please, creating a “mash-up” of several lives or a “life-mix” (Turkle 2011: 162), allows us to vary the degree of coherence in our self-understanding—we can consciously split our personal identificatory narrative into two or more, living parallel lives guided by incompatible personal narratives. The switch from the unified self-understanding to a view of the self as a collection of distributed agents that are fairly independent of each other is a well-known topic in psycho-analytic theory, especially in “ego-state theory.”⁹ As Turkle points out, there are therapeutic benefits of such daily “cycling through online personae” (2011: 134) since “when identity is multiple in this way, people feel ‘whole’ not because they are *one*, but because relationships between aspects of self are fluid and undefensive. We feel ‘ourselves’, if we move easily among our many aspects of self” (ibid. 194). Interestingly, in the quoted passage Turkle does not speak of “multiple identities” but of an “identity [that] is multiple” and the multiplicity in question pertains to “aspects of the self,” which is a more moderate interpretation. Ontologically speaking, the new cultural practice of operating several personal narratives simultaneously amount to engenders structural transformations of the empirical Ego—how we ourselves understand who we are; these transformations seem to lie on a gradient scale between, on the one hand, “weak disintegration” into a multiplicity of compatible personal narratives and, on the other hand, “strong disintegration”, which involves a transition from static unity to dynamic unity, where internal coherence or the compatibility of parts is no longer required.

⁸ Back et al. (2010: 372) report that “more than 700 million people worldwide now have profiles on on-line social networking sites”; currently there are 800 million users on Facebook alone and about 4 billion users on listed social networks, http://en.wikipedia.org/wiki/List_of_social_networking_websites (accessed 5 Dec 2011). The proportion of global internet users to users of social networks seems to be about 1: 4, http://en.wikipedia.org/wiki/Global_Internet_usage, accessed 5 Dec 2011.

⁹ The specific link between multiple online identities and “Ego State Theory” (developed by Paul Federn in 1952, extended by John and Helen Watkins (1992)) is not, to our knowledge, to be found in the literature yet, but would seem to be empirically significant for the latter.

In terms of our formal notation above, our new cultural capacity of living a “mash-up” of several lives, real and virtual, engenders a spectrum of structural transformations that can be characterized by the following extremal points¹⁰:

Weak disintegration Personal identity remains a function that selects, for a temporal interval, individual capacities to form a personal narrative that represents who we understand ourselves to be, but now the value of the selection function at any interval is a set of non-overlapping personal narratives: i.e., $f_{pi}: T_{a1} \times C_{a1} \rightarrow V$, with $f_{pi}(t_i, C_{a1}) = S_i$, where $S_i = \{V_{a1,ii}\} \in V$. If the disintegration is moderate the set S_i is consistent and all personal narratives operated at an interval are compatible with each other and can be consciously entertained at the same time.

Strong disintegration The selection function of personal identity yields at any interval not a set but a sequence of incompatible personal narratives; the length and order in which these different personal narratives are “cycled through” or consciously entertained is variable. $f_{pi}(t_i, C_{a1}) = S_i$, where $S_i = \langle V_{a1,ii} \rangle \in V$.¹¹

Note that both weak and strong disintegration amount to structural changes in the values of a single selection function of personal identity; in other words, neither weak nor strong disintegration involve the pathological condition of “multiple personality disorder,” which in the present framework can be modeled as an individual’s having several selection functions of personal identity. The typical sign of pathological disintegration consists in the dissociation of entire streams of experience, the experiential subjects of which are not in contact with each other.¹² In contrast, in non-pathological disintegration, there is a single stream of experience (here: set of individual capacities C_{a1}) and a single experiential subject who remains in control of the operation (even though not of the construction) of the multiple personal narratives selected at any interval.¹³ Thus, while the disintegration of a unique personal identificatory narrative into a multiplicity of such narratives surely is a structural change that can be attributed to the “disembodied internet,” it does not, as such, imply that any of the many persons we take ourselves to be has reduced ethical sensitivities or is less of a moral subject.

The second candidate for a structural change of identity constructions effected by the ‘disembodied internet’ is what one might call the “delayed phenomenology” of feelings. This is a structural change not in the sense that it partitions the set of capacities or experiences that constitute a person’s identificatory narrative, but in the sense that it changes the conditions under which certain capacities or experiences are formed. Turkle cites teenagers who claim that they first send verbal descriptions of their emotions to their friends on social media, and only

¹⁰ These transformations may be classified more finely; cf. for example (Stutzman and Hartzog 2009), who sort a “continuum of boundary regulation behaviors emerging from multiple profile maintenance” according to main underlying motives.

¹¹ Here, we omit a precise definition of the variability of the order of the sequence.

¹² See, e.g., Watkins and Watkins (1992).

¹³ In the nonpathological case, the relationship between the different empirical egos as defined by the different personal narratives selected for an interval is one of fore-grounding and back-grounding at different degrees (Watkins and Watkins 1992). Sicart’s (2009:70ff) phenomenological description of the relationship between player subject and game subject in online games displays striking similarities.

subsequently, depending on social approval or disapproval, feel the emotions described, as if “discovering their feelings by texting them” (Turkle 2011: 198). To the extent that these claims can be taken to be accurate introspective reports on the onset of a feeling, one could argue that the mobile internet turns inner episodes that we used to experience as “Cartesian”: i.e., immediate, private, and incorrigible, into a new sort of inner episode with delayed and socially conditioned phenomenology. As with the multiplication of personal narratives we considered above, we find that this kind of structural change in identity constructions effected by the “disembodied internet” does not necessarily imply a reduction of ethical sensitivities. However, while the disintegration of personal identificatory narratives is meta-ethically neutral, this kind of structural change would imply an impairment of ethical capacities if one were committed to certain meta-ethical positions, e.g., a non-rationalist theory of moral motivation in the tradition of Hume or the ethics of care.

Our third candidate for structural change on identity constructions effected by the “disembodied internet” is a praxis we call “temporary depersonalizations”. To draw again on one of Turkle’s (2011) observations, in the course of mobile internet communication we have come to accept different social conventions concerning the perception of presence and absence. In the middle of informal private meetings we permit “absences” of physically present people who, tending their smart phones, intermittently vanish into the virtual environments they carry with them (Turkle 2011: 154). When our interlocutors become “absent” in this sense, abandoning their communicative duties towards us, we typically cognitively “tune out” and busy ourselves with a practical task. In this way, we create a cognitive scenario that lets the “absent” interlocutor recede into the background of an instrumental praxis, relieving him or her temporarily of any social duties as interlocutor. This move can be more technically described in Heideggerian terms (Heidegger 1927: §18) as a switch in our interactive relationship with the interlocutor from “Mitsein” (“Being-with”, the social dimension of existence) to “Zuhandenheit” (“ready-to-handness”, instrumental dimension). We interact with the “absent” interlocutor no longer as an item in the social sphere but rather as “Zeug” (“equipment”), i.e., as an item that belongs into the context of a practical task. Equipment may be in the foreground or background relative to the phase of the practical task. If a piece of equipment is in the foreground—e.g., the screw driver you are using now in order to open the outlet—it has “inconspicuous” presence in its functional role; if it is in the background—e.g., the screwdriver you have put down in order to use the pliers and cut the wire—it has inconspicuous absence. As piece of equipment in the background, the interlocutor is inconspicuously absent in ways that defuses potential tensions—the interaction is taken out of the context of social obligations and embarrassments yet a context of meaningful engagement is maintained. As in the previous case of structural change, the cognitive switch involved alters not just the content of what we experience but also the way we experience and form our personal identificatory narratives. The modification is as such ethically neutral—we tend to consider it as an act of social etiquette. On the other hand, one might worry that the frequency and ease with which we perform such temporary depersonalizations has negative effects on our capacity for moral agency. But again, as in the previous case of structural change, we need to add

the premises of a suitable theory of moral motivation in order to make the case that temporary depersonalizations amount to ethical degradation.¹⁴

In sum, then, in this section we clarified several senses of ‘identity’ and considered three ways in which the “disembodied internet”—i.e., the internet as communication with purely semantic representation of the interlocutors—structurally affects the construction of personal identificatory narratives: the multiplication of personal identificatory narratives, the social conditioning of feelings, and the temporary depersonalization of interlocutors. We drew attention to the fact that none of these three structural modifications by itself implies an impairment of our capacity of moral agency.¹⁵ In order to establish such a consequence, we need to add a theory of moral motivation or moral cognition that would allow us to explain why disembodied communication diminishes our ethical sensitivities. Thus we turn now to the question of why *embodied* communication can be assigned a special role for moral cognition and the formation of a moral self.

2 Dialogue and Cognitive Orientation

In the early stages of the internet, it was noted that communicating predominantly by a “disembodied medium” carried some risks: the emotional poverty of internet communication was found to diminish semantic effectiveness in contexts of personal communication and to affect psychological well-being (cf., e.g., Guye-Vuilleme et al. 1999; Kraut et al. 1998). Only more recently, however, it is becoming clearer how observations about characteristic consequences of disembodied communication might fit into a more comprehensive view of the connections between embodiment, agency, and cognition. During the last decade (mainstream “analytical”) philosophy of mind and cognitive science increasingly have turned towards a new paradigm of research that is approached under various labels: “embodied cognition,” “embodied mind,” “enactionism,” or “interactivism.”¹⁶ Notwithstanding differences in method and disciplinary focus, in their basic theoretical assumptions these research programs share a common line that can be summarized into the following two principles: (1) cognition is essentially dependent (both in terms of diachronic and synchronic dependency) on the sensorimotoric interaction with the physical environment of the cognizant agent; (2) in particular, conceptual cognition, including conceptual experience, depends on pre-conceptual cognition that cannot be exhaustively described in terms of categories that apply to the processes “within” the cognitive subject, such as the classical Cartesian notion of a feeling. Embodied cognition research addresses

¹⁴ Ling and McEwen (2010) similarly describe decisions about the “parking” of personal conversations as involving “two levels of reflexivity,” i.e., as the embedding of social etiquette in an ethical frame.

¹⁵ The three modifications considered are those currently highlighted in the literature; we do not exclude that there may be some other structural modification of our sense of self that would by itself imply ethical degradation. To restate, we merely wish to point out that—in so far as it turns on these three modifications—the debate about whether the disembodied internet affects our capacity of moral agency is based on implicit assumptions about the role of embodied communication for moral motivation, and that the explication of these assumptions can be theoretically and technologically fruitful.

¹⁶ See Calvo and Gomila (2008).

fundamental questions of cognition, such as the interactive modification of qualia contents, the interactive emergence of representation, and also increasingly the interactive foundations of social cognition and empathy.¹⁷ However, specifically the connections between embodied communication and the formation of our moral sense of self have not yet, it appears, received attention. As we will try to show in this section, by reframing the phenomenological description of dialogue within the theory of “cognitive orientation”, an approach in cognitive psychology that fits with basic methodological tenets of embodied cognition, we can formulate an empirical research hypothesis about the role of embodied communication for the formation of a moral sense of self. We argue that this hypothesis has sufficient initial plausibility to merit proper empirical investigation. Once it is in clearer view which cognitive functions one may plausibly assume to be at work in embodied communication, it is possible to discuss the prospects of specific proposals for “embodying” the internet, which we undertake in Section 3.

Let us begin, then, with a look at the phenomenological study of direct personal communication. While Merleau-Ponty’s phenomenological studies of the role of action for perception are frequently referred to in studies in embodied cognition.¹⁸ Martin Buber’s inquiries in the phenomenology of dialogue and non-alienated self-relationships so far are not yet relevantly linked to the debate. The systematic aims of Buber’s poetic texts go far beyond our issue here, and we will take the liberty of setting aside much of the wider philosophical (and theological) context of his observations, while translating the latter into the more prosaic idioms of contemporary philosophical analysis in order to facilitate interdisciplinary interfacing.

Buber invites us to reflect on the peculiar differences between two cognitive modes. The first mode we are in, typically, when we describe our surroundings in terms of classificatory predicates. In this mode of cognition, which he calls “analytical experience” (“Erfahrung”), we apply concepts and predicate, i.e., we perform classificatory judgments. In performing the cognitive act of a classification, judgments we relate to something as a classifiable item, an “object”—in Buber’s terminology, we stand to something in an “I-it” (“ich-es”) relationship. In contrast, the second mode of cognition does not produce any classifiable objects but remains inquiring attention: a dynamic state of cognitive exploration. Typically, we are predominantly in the second mode of cognition when we are engaged in a dialogue with another person. In direct personal communication, Buber claims, we do not classify our interlocutor; rather, the “I” is arrested in an *encounter*, is tied up with “what lies across”: a “Thou”, and is caught up in an “engagement” (“Beziehung”) that is felt as an ongoing search for meaning. In essence, the cognitive mode of the “I-Thou” relationship is characterized by four key factors. The information obtained is (a) *immediate*, i.e., processed without conceptual judgments and (b) *holistic*, i.e., not focused on single feature; the cognitive posture in “I-Thou” relationships creates (c) a feeling of being “called upon,” i.e., of being subject to implicit norms of *communicative* interaction, but also (d) the

¹⁷ See, e.g., O’Regan and Noë (2001), Bickhard (2003); Meteyard et al (in press); on social cognition see the extensive literature cited in Gallagher (2008) and Reddy (2008).

¹⁸ Cf., e.g., Varela et al. (1991), Clark (1997), Petitot et al. (1999), Dreyfus (2001), Gallagher (2003), and Noë (2004).

feeling of *generating* oneself in the course of one's reaction: "I become upon [i.e., through engagement with] the thou; becoming me I am saying thou."¹⁹

It is not difficult, we trust, introspectively to confirm these phenomenological claims—we typically do not "observe" another person but "take in"²⁰ what is presented to us, without reflective discrimination between situation and person, and take in the person as a whole, getting a "feel" for the person and reacting "intuitively" (as we are often wont to put it), guided by clues rather than by conscious evaluation and decision-making. However, it is crucial to realize that while the "I-thou" mode of cognition has its prototypical occurrence in dialogue, in Buber's view this mode is not restricted to personal encounters—everything *can* be approached in the mode of intuitive cognitive engagement we are most familiar with from our concrete encounters with other people. Vice versa, the "I-it" mode of cognition is not limited to our interactions with objects: we also *observe* people, we *plan* our reactions carefully and consciously, etc. In fact, in most situations we switch back and forth between the two modes of "intuitive" taking in versus classificatory cognitive processing.²¹

As we shall try to make plausible now, there are striking and useful connections between (a) the mode of cognition that Buber identified as the "I-thou relationship" and (b) the process of "cognitive orientation," the development of agentive motivations as described in cognitive psychology. In particular, this link will allow us to make good sense of Buber's idea that the constitution of a self happens in the course of a mode of cognition that is not unique to, but has its primary occurrence in, dialogue.

To begin with, let us note that the "I-thou" mode and the "I-it" modes are what one might call 'functionally ordered': it is our "fate" that "every *thou* in our world must become an *it*" (Buber 1923/1979:24). In other words, human cognition is in Buber's view a process in the course of which contents that are apprehended via "intuitive, immediate, comprehensive taking in" are subsequently made explicit by conceptual articulation into distinct classificatory semantic units, which in turn are operated on in routines of inference and learning. While single perceptual acts have thus, phenomenologically speaking, two phases, as generic cognitive activities the two modes of cognition occur simultaneously, and each may be in the foreground or background of introspective awareness. The functional ordering of explication—an innate tendency to transform and partition what we take in into distinct conceptual units—does imply a temporal ordering for any single content that is getting conceptualized, but it does not imply that the "I-thou" mode and the "I-it" mode are sequestered into different time intervals; rather, they are two co-occurrent modes of cognition. In a sense, the basic interpretational model at work in Buber's description of cognition is still the Kantian account: a streaming of sensory information is gradually organized by spatial and temporal ordering relations ("forms of intuition") and other relational templates

¹⁹ Cf. Buber (1923/1979 (10th ed.): 18) and Lang (1963). Besides observing that we feel "called upon" by norms, Buber laconically elucidates the remaining aspects as follows: "Die Beziehung zum Du ist unmittelbar"; "Was weiss man vom Du?—Nur alles. Denn man weiss von ihm nichts Einzelnes mehr"; "Ich werde am Du; Ich werdend spreche ich Du."—The first two elements of Buber's distinction between two modes of cognition seem to link up with (Dreyfus 2001) observations that the 'disembodied internet' serves well to teach declarative knowledge but is ill-suited for the imparting of skills, which apparently requires that the learner is in the embodied co-presence with someone who is exercising these skills.

²⁰ Buber (1932) contrasts "beobachten" (observing) with "innewerden" (getting-changed-by-taking-in).

²¹ Essentially, Buber introduces the "I-it" mode and the "I-thou" mode essentially as two forms of intentionality: two ways of being-there as a human being in interaction with the world.

(“schematization” and “categorization”) that cluster information into the units we consciously experience as material things with features. Unlike Kant, however, Buber does not consider the preconceptual phase of cognition as one where information is grafted onto a passive recipient; rather, he presents it as the subconscious engagement of a subject that is tied to the environment by a “mutuality of giving” in the style of an entangled bootstrapping process researchers in embodied cognition call “structural coupling”. Again, unlike Kant, Buber takes the preconceptual phase of processing to be epistemically accessible to us (though not, of course, by conceptual experience). In fact, in Buber’s view, it is only in that dynamic state of coupled interaction that we are truly *present* and are present to ourselves, though not as conceptualized “selves” yet—in the coupled ‘state’ of the ‘I-thou’ mode of cognition we are aware of ourselves merely as activity that enables contrast and dynamics of an interaction: as “Gegenwartende and Gegenwährende” (“a waiting towards and giving room”; (Buber 1979; *Ich und Du* & here quoted after the 10th edition 1979):19f).

It is in particular this cognitive posture of being aware of oneself as “waiting towards and giving room” that Buber takes to be crucial for the formation of a moral self: the state of being “attentively arrested”, of explorative listening to a situation that has not yet been partitioned by the cognitive prejudice of our classificatory judgments. If we were to express this claim in terms our identity model in Section 1, Buber suggests that the formation of a moral self hinges on the extent to which the generation of our identificatory narratives can remain open to the “demands of the situation” without rushing into a quick and mechanic classification. To characterize, as we have done above, this formation process as a function, a mapping of elements, may be a suitable formal metaphor for the “I-it” mode of cognition but is potentially misleading for the purpose of conveying the complex dynamics of “structural coupling” in the “I-thou” mode.²²

How to relate Buber’s poetic descriptor of “waiting towards and giving room” to a suitable theoretical idiom in terms of which one could empirically investigate how the “I-thou” mode of cognition features in moral motivation or moral cognition? Interestingly, the phenomenology associated with this cognitive posture of “taking in while waiting towards and giving room,” namely, a feeling of attentiveness, exposure, and of being outside of any instrumental routines, we know not only from personal encounters in dialogue but also from situations where we try to *orientate ourselves* in physical or figurative space. Recall what it is like to leave a building on your university campus from an exit you have never used before: for a moment you hesitate, i.e., your practical routines are interrupted, you are slightly more alert, sensing exposure, and then you have got our bearings again and walk off. Typically, this process goes all but unnoticed, since it is very brief and does not involve any conscious self-positioning on a memorized spatial map. But in more unfamiliar terrain orientation can become “pronounced,” turning into a conscious and reflected complex cognitive interaction with the environment where we go back and forth between some immediate preconceptual positioning and conscious classificatory

²² Recent empirical support for Buber’s claim that the ‘I-thou’ mode involves a relational self comes, for example, from cognitive science and developmental psychology (Nagy 2009) documenting the capacities of neonates and young infants (less than 2 years old) to engage in social relations and have self-conscious experiences; this has been taken to suggest that besides our conceptualized self there is an experienced self that is “innately relational,” cf. (Reddy 2008), (Schilbach et al. 2006).

experience of directions and landmarks. In Buber’s terminology, such pronounced orientation consists of a continuous mutual backgrounding and foregrounding of the “I-thou” mode and the “I-it” mode of cognition.

That “orientation” is a mode of cognitive processing that is going on ahead of and alongside conceptual experience and inference was already recognized by Kant—who called it a “felt need of reason” (ein “gefühltes Bedürfnis der Vernunft”)—and has been further explored within contemporary German philosophy, primarily with a focus on what we called above “pronounced orientation”, i.e., a conscious and reflected activity of getting one’s bearings within not just physical space, but any space: social space, logical space, or value space.²³ The new category of a “structural coupling” between organism and environment seems well-suited to capture the dynamic state of interactivity at issue: orientation is the structural transformation of both organism and perceived environment, an entangled bootstrapping of orientatedness in the organism and the differentiation of the environment into orientating cues (the coming about of affordances).²⁴ Already in 1976, psychologists Hans and Shulamith Kreitler argued—although not in these terms—that orientation is a form of structural coupling. In a work that from today’s point of view has almost uncanny visionary qualities, Kreitler and Kreitler present a “theory of cognitive orientation” which takes orientation to be the model of cognition in general, anticipating many of the guiding principles of contemporary theories of embodied cognition. Here we can only sketch some of core ideas of the approach, which in the present context merely serves the purpose of relating Buber’s insights to more contemporary idioms of research.

The theory’s heuristic point of departure are Pavlov’s studies on the so-called “orientating reflex” in higher animals, which consists in a “matrix of specific somatic, automatic, electroencephalographic and sensory reactions” and two behavioral aspects: the organism’s directing itself towards the stimulus for increased sensory reception and active exploration of the stimulus in the physiological state of attentiveness (higher muscle tonus, lower perceptual thresholds) to determine which further course of action: fight, flight, or feed, is appropriate. However, as Kreitler and Kreitler emphasize, the “meaning action” in the orientating ‘reflex’ is not passive classification but active exploration. Orientation is, in the first instance, the interactivity of organism and environment in the course of which a network of (preconceptual) significances is (“online”) established and continuously revised.²⁵ Repetition of such orientational

²³ In the course of discussing the problem of the difference between left and right, Kant came to identify orientation as a type of cognitive processing *sui generis*, especially in the short piece (Kant 1786) on “What does it mean to orientate oneself in thought?”. For recent work on the systematic significance of “orientation” for epistemology, philosophy of mind, action theory, and ethics compare in particular the essays collected in Stegmaier (2005).

²⁴ The notion of structural coupling (see e.g., Maturana and Varela (1987):75) was introduced by Maturana and Varela in the late 1970s; within the contemporary discussion about embodied cognition the three components of the notion: coordination, coevolution, and coupling (in the sense of the dynamic entanglement of two systems due to interlocking processes) have received different emphasis by different authors.

²⁵ Cf.: “The placement of meaning action at the core of perception implies that there is no one particular stage or fixed point [...] at which the selection of what to attend to is established” (Kreitler and Kreitler 1976: 61). Kreitler and Kreitler also promote the radical thesis of embodied cognition that this interactivity is all there is to cognition, with semantic ‘units’ being formed as the emergent quasi-objects. See (Seibt 2005). Thus—and this seems to have gone unnoticed by proponents of embodied cognition—the theory of cognitive orientation is the place where a so-called “pure theory of embodied cognition” (Calvo and Gomila 2008: 17) has been formulated for the first time.

interactivity generates habitual routines in the ongoing orientating interpretation of the agentive situation, and it is these routines that we go through in common classificatory experience.

In the interactivity between organism and environment preconceptual significances are combined into “cognitive orientation clusters” which provide for the organism’s orientatedness relative to different ‘spaces’, including—in the case of humans—normative and axiological spaces. In a nutshell, the “theory of cognitive orientation” presents a view of cognition as an interactivity of a coupled organism–environment system that continuously generates, based on interpretational routines, meaning units from preconceptual significances, thereby positioning the organism within a matrix of factual and normative valuations, which constitute the agent’s motivational state.

Kreitler and Kreitler’s theory of cognitive orientation has considerable empirical support and therapeutic applications.²⁶ In the present context we use it, to restate, mainly for heuristic purposes to reframe and actualize Buber’s phenomenological description of our mode of cognition in personal encounters. The feeling that accompanies the “I-thou” relation: exposure, immediacy, alertness, and openness, is the phenomenal evidence for a mode of cognition that we know otherwise from occasions of orientation, we pointed out, and drew attention to that there is, in fact, a well-worked theory of cognition that accords well with Buber’s observations about the “I-thou” relation being the basic mode of cognition, as well as the relationship between the “I-thou” and “I-it”. In a contemporary idiom that lends itself more easily to theoretical interfacing with empirical research in cognitive science and robotics, Buber’s observations amount to the claim that cognition is the interactivity of the coupled organism–environment system during which preconceptual significances and orientational environmental cues are jointly bootstrapped, and that this basic interpretatory interactivity leads—perhaps via Hebbian learning—to the formation of semantic units as the emergent stable routines of this dynamics. Moreover, the theory of cognitive orientation also supports Buber’s claims that cognition in the mode of the “I-thou” is valuative: cognitive orientation generates semantic units that are pragmatically charged, thus determining even in classificatory experience the agent’s motivational “state”. Most importantly for our purposes here, the interactivity of cognitive orientation is a “positioning” of the organism relative to a number of valuative spaces in the course of which a self-conception or personal narrative is generated. Thus, if we equate the interactivity of cognition that Kreitler and Kreitler label cognitive orientation with Buber’s “I-thou” mode of cognition, it is indeed in the “I-thou” mode of cognition that we *are* ourselves—a self is the dynamics of cognitive orientation in the course of which conceptions about the world and our self-conceptions or personal narratives are generated.

Finally, even though here a closer discussion of the empirical research on the theory of cognitive orientation would be necessary, it is also plausible, we submit, that the latter can be aligned with two further phenomenological claims of Buber’s. First, even though the “I-thou” mode of cognition is domain-independent, dialogue,

²⁶ The theory has been empirically investigated in over 60 studies in social and educational psychology; therapeutic applications pertain to motivational ‘reprogramming’ in cases of substance abuse. See, e.g., Kreitler and Kreitler (1986a, b, 1987, 1990, 2004) and Kreitler (2001, 2002).

i.e., embodied personal communication, in particular ways opens us to this mode, and arrests us in it. In contrast, things—especially artifacts like cups and computers—are fitted into functional contexts of instrumental use and reinforce the “quick and mechanic” interpretational routines of the interactivity of orientation in classificatory experience. The person we encounter in direct communication, on the other hand, resists such instrumental classifications, requiring continuous—however unconscious—reorientation about “what she or he means for me” at the preconceptual level; this prolonged “coupling” or *intensified orientation* creates at the phenomenal level the peculiar “entanglement” we feel with another human being in direct communicative encounter.²⁷ Second, by arresting us in the interactivity of intensified orientation, direct personal communication or dialogue prolongs the phase during which a positioning in valuative and normative spaces, among others, is updated. Thus dialogue increases the chance for the formation of a motivational state in the agent that is open to the moral demands of precisely this situation, and, more generally, increasing the time during which agents are prevented from classificatory prejudice and remain in a state of dynamic attentiveness to all valuative dimensions, including moral values.

Let us retrace our steps and summarize the results of this section. We entered this section with the question of whether disembodied communication threatens our sense of self as the very ground for a coherent and morally charged construction of personal identity. We suggested that this question can be fruitfully approached by taking a closer look at Buber’s phenomenological analysis of dialogue, a prominent place in philosophy where direct embodied personal communication is systematically tied to the reality of the self and the interlocutors’ capacity for moral judgment. Reframing Buber’s observations about the “I-thou” mode of cognition within the theory of cognitive orientation and linking both to relevant terminology of the embodied cognition program in cognitive science, we have now arrived at the thesis that embodied communication among persons arrests the interlocutors in a mode of cognition—the interactivity of orientation in coupled systems—that as such is ubiquitous, but is commonly preempted, or placed into the background, by the classificatory judgments of conceptual experience. This abridges the role that the interactivity of orientation can play in the formation of the agent’s motivational state. In contrast by prolonging the periods in which human agents are engaged in the interactivity of orientation, direct encounters enable the interlocutors to feel that interactivity as the primary reality of their selves, as that which engenders the construction of personal narratives. Phenomenologically speaking, in dialogue or direct embodied personal communication we have a deep sense of self without knowing who we are. At the same time, by prolonging periods of interactive intensified orientation, dialogue forces agents to remain in a dynamic state of moral

²⁷ We distinguish, at the phenomenological level, three forms in which we are aware of being in the “I-thou” mode of cognition. First, and that is the default, the “I-thou” mode or the preconceptual phase of cognition is in the background; second, it may be “intensified,” i.e., in the foreground and something we are aware of; third, it may be “pronounced”, i.e., something we are reflectively aware of. Orientation in physical or figurative space is always intensified and frequently also pronounced, especially when we know we have lost our way, while dialogical orientation is typically merely intensified and becomes pronounced only when we try to “figure someone out.”

attentiveness, allowing for valuative repositionings relative to various normative “spaces”.

3 Telenoids as ‘Thou’s?

The results reached in the previous section, we believe, can be put to good heuristic use both in empirical research and in the philosophical discussion of proposals for the embodiment of internet communication. We shall here illustrate the former part of this claim by taking a closer look at a recent development in telepresence technology, namely, the use of teleoperated robots in remote communication. While some approaches to telepresence technologies—e.g., devices developed by CISCO (HSL 2007) and Telstra (Meredith 2008)—focus on the sensory enrichment of on-screen or hologram presentations of interlocutors, teleoperated robotics aims to provide some form of embodied presence of one of the interlocutors (hereafter: “the operator”) in the physical location of the other interlocutor (hereafter: “the interactor”). The best-known implementation of this strategy might be the *Geminoid HI* and *Geminoid F* robot, developed by Hiroshi Ishiguro at the ATR Intelligent Robotics and Communications Laboratory in Kyoto, Japan (for details, see Nishio et al. 2007); these robots replicate the physical features of one specific interlocutor in great detail. More recently, however, Ishiguro changed his design strategy and developed the *Telenoid*, a teleoperated communication robot that displays the features of a “minimal human” (Ogawa et al. 2011, p. 2), i.e., a shape that can count as the least common denominator of human physical features (see Fig. 1).

The lack of “detail”, so the assumptions of this design strategy, will allow the “interactor” (person in the physical presence of the Telenoid) mentally to “dress” the robot with the physical details of the robot’s operator. For example, housebound grandparents around the world would be able mentally to project their grandchildren’s physical features onto the robot’s body, facilitated by a physical appearance that is neutral with respect to gender, age, and ethnic phenotype. In the current setup, the robot’s interactive capabilities are as follows: the robot’s operator at the remote site uses a computer, the internet, a webcam, a microphone, and special teleoperation software; the computer captures voice and tracks the operator’s head movements, and transmits some of the neck movements to the Telenoid; in addition, the operator can also push buttons to activate other behaviors, such as a hugging motion. The operator

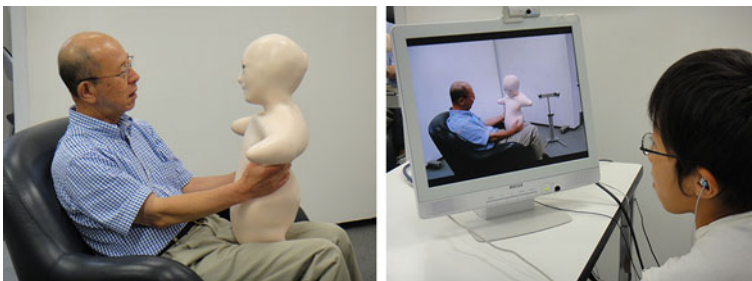


Fig. 1 The Telenoid robot. Photos by ATR Intelligent Robotics and Communication Laboratory

receives his sensory input from the encounter between the Telenoid and the interactor on the computer screen and through the headphones.

A primary domain of the intended application of Telenoids is elderly care, which has raised factual and ethical questions. To what extent can and may Telenoid communication replace direct dialogue with family and friends? To what extent can and may internet communication via Telenoids replace direct dialogue with those who care for us as health counselors or teachers? In discussing the prospects of social robotics, some authors have already decided that any attempts of replacing human-human interactions by human-robot interactions are not only hopeless but also morally wrong. Thus Dreyfus (2001) categorically rejected the idea that robots could be suitable stand-ins for people: “Even the most gentle person-robot interaction would never be a caress nor could one successfully use a delicately controlled and sensitive robot arm to give one’s kid a hug. Whatever hugs do for people, I’m quite sure telehugs won’t do it” (2001: 69). Similarly, commenting on a pilot study on Telenoid communication in March 2011, the Chairman of the Danish Governmental Ethical Council feared that “relational technology could result in blunting human emotionality,” echoing the more general pronouncement of the Danish Ethical Council that the employment of social robots is ethically problematic.²⁸

Such evaluations proceed from the presupposition that Telenoids are (a) to replace the presence of an individual human being (the operator), or (b) to replace the presence of some human being, i.e., to replace human personal encounter. However, as we want to argue here, these presuppositions are by no means inescapable. Approaching matters more cautiously, we note that the primary function of the Telenoid is to introduce certain aspects of bodily communication into internet communication. This is distinctly different from the task of embodying a human being, let alone an individual person. In order to discuss the potential benefits or detriments of employing teleoperated communication robots, we need to investigate first whether and in which sense the latter fulfill their primary function in (partially) embodying internet communication; in a second step, then, one might ask whether the relevant form of embodied internet communication will hold out the prospect of providing remedies to those problematic features of disembodied internet communication that are said to threaten our sense of self as the very ground for a coherent and morally charged construction of personal identity. Our following remarks are intended as contributions to more relevantly focused questions of this kind.

One of us, Marco Nørskov, participated in the aforementioned recent pilot testing of the Telenoid in care centers and homes of elderly people in the Svendborg municipality, Denmark, in March 2011. As Nørskov has argued elsewhere (Nørskov 2011) based on his experiences in this pilot study, teleoperated communication robots call for a new approach to the classification of phenomenological human-technology relations as so far codified in D. Ihde’s familiar fourfold scheme (Ihde 2004). For example, whether the relationship between operator and Telenoid appears as—in Ihde’s senses of these terms—“embodiment relation” or “hermeneutic relation” depends, as Nørskov observes, not only on the dexterity of

²⁸ See *Information*, March 25, 2011 and Ethical Council (2010).

the operator but also on the operator's personality and communicational style (e.g., whether the operator is someone who easily "loses" himself in the emotional atmosphere of the conversation, thus forgetting about the 'hermeneutical distance' to buttons that administer hugs, etc.) and on the particularities of the interlocutor (e.g., if the interlocutor is a good friend, robot-mediated hugs feel "embodied" in Ihde's sense, otherwise the "hermeneutic relation" prevails). The relationship between robot and interactor indicated similar variations in degree and kind, depending on the personality of the interactor (e.g., not only on the extent to which the interactor would give himself or herself into the emotionality of the situation but also on the interactor's emotional sensitivity, the extent to which the communication situation and restricted expressiveness of the Telenoid were felt as inhibiting). Most important for our present question, however, is Nørskov's observation that many of interactors—even among those who did not understand that Telenoids were teleoperated—refrained from treating the Telenoid as *any* of the familiar types of entities: dolls, pets, or people. Interactors attended to the Telenoid with curiosity and engaged in orientational exploration by physical approach, handling, and communication. (The various phases of such orientational exploration can be watched on YouTube, <http://www.youtube.com/watch?v=bvfVznGU2CE>, featuring people interacting with the Telenoid at the 2010 ARS Electronica in Linz, Austria).

Harking back to our considerations in the previous section, it appears that a Telenoid has the potential to arrest interactors in the interactivity of orientation or cognition in the mode of the "I-thou" relation, even though it is decidedly *not* understood as a human "thou." Significantly, internet users who read about the Telenoid or watched the mentioned YouTube clipping reacted with strong expressions of the feeling of 'uncanniness' ("I shudder", "scary"; "super creepy"; "freaks me out" etc.), which is a common response to the unfamiliarity of the familiar. However, real-life interactors did not, with a few exceptions, express similar aversions after the encounter with the robot, which could be attributed to the fact that in the physical presence of the robot the latter's unfamiliarity was transformed by the physical coupling effected in the cognitive dynamics of intensified orientational interactivity. Even though it seems that the Telenoid is not experienced as a replacement of the operator, nor as a replacement of a human being in general, there are some first indications that it can put interactors into a mode of cognition that resembles the "I-thou" mode of cognition in human dialogue, with its typical prolonged phases of orientational attentiveness, the feeling of exposure, of immediacy, of holistic taking in. For example, the interactors predominantly displayed positive attention and continued to do so even if they verbally expressed a feeling of cognitive dissonance in view of something for which they had no classificatory routines. The knowledge that the Telenoid was teleoperated did not seem to affect the mode of interaction: while such knowledge added to the expression of cognitive dissonance, it did not increase or decrease positive attention to the robot. These first impressions have heuristic value, we believe, for the design of future empirical research in "second person cognitive science" (Nagy 2009). The primary task of such empirical studies could be to investigate whether the orientational attentiveness towards a Telenoid indeed can qualify, at the neurological level, as the "I-thou" mode of cognition that interactors are in when they are engaged in personal dialogue. The next task would be

to explore foundations of moral cognition and to test the phenomenological claim presented above, namely, that the prolonged withholding of interpretational instrumental routines allow for a “valuative updating” and provides us with a sense of self as the activity that generates our personal narratives and opens us to the ethical values of the situation.

4 Conclusion

Internet communication has been said to facilitate constructions of personal identificatory narratives (of how we understand ourselves at a certain time) that are not sufficiently informed by ethical values, since they are no longer (exclusively, or even only predominantly) constrained by our interactive experience in direct personal communication. In order to investigate whether new technologies for the embodiment of internet communication can compensate for this deficit, we need to understand why direct personal communication engenders ethically sensitive constructions of self-understanding. We suggested in this paper that the key to such understanding lies in a closer empirical investigation of the role of a preconceptual cognitive interactivity that we experience consciously when we orientate ourselves. Based on phenomenological observations by Martin Buber and the “theory of cognitive orientation”, an early theory of embodied cognition, we formulated the threefold empirical hypothesis that (a) cognitive orientation is a cognitive interactivity during which we are sensitive to the valuative dimension of a communicative situation; (b) that the interactivity of cognitive orientation leads from a preconceptual experience of “self” to the generation of personal identificatory narratives in terms of which we understand ourselves in classificatory predicates; and (c) that direct personal encounter prolongs the phase of cognitive orientation, thus increasing the chance that our personal identificatory narratives are informed by the ethical values of the situation. We presented some preliminary indications that teleoperated communication robots—the so-called “Telenoid” developed by H. Ishiguro—also can prolong the phase of cognitive orientation in at least one of the communication partners. Altogether, the aim of this paper has been to draw attention to the significance of empirical research on cognitive orientation as a special interactive mode of cognition grounding value experience. Whether and which new technologies for embodying the internet will facilitate ethically sensitive constructions of personal identificatory narratives will depend on the results of the empirical investigation of the threefold hypothesis we have presented here.

References

- Back, M. D., Stopfer, J. M., Vazire, S., Gaddis, S., Schmukle, S. C., Egloff, B., et al. (2010). Facebook profiles represent actual personality, not self-idealization. *Psychological Science*, 21(3), 372–374.
- Bickhard, M. H. (2003). Process and emergence: normative function and representation. In J. Seibt (Ed.), *Process theories: crossdisciplinary studies in dynamic Categories* (pp. 121–155). Dordrecht: Kluwer.
- Bickhard, M. H. (2010). Interactive knowing: the metaphysics of intentionality. In R. Poli & J. Seibt (Eds.), *Theory and applications of ontology. Philosophical perspectives* (pp. 207–229). New York: Springer.
- Buber, M. (1932). *Zwiesprache*. Berlin: Schocken.

- Buber, M. (1979). *Ich und Du*. 10th edition, first appeared in 1923. Heidelberg: Lambert Verlag.
- Calvo, P., & Gomila, A. (Eds.). (2008). *Handbook of cognitive science: an embodied approach*. London: Elsevier.
- Clark, A. (1997). *Being there: putting brain, body, and world together again*. Cambridge: MIT Press.
- Cocking, D. (2010). Plural selves and relational identity. In J. Van den Hoven & J. Weckert (Eds.), *Information technology and moral philosophy* (pp. 123–141). Cambridge: Cambridge University Press.
- Csepeli, G. (2009). Masquerade in the Blogroom. In K. Nyiri (Ed.), *Engagement and exposure* (pp. 112–119). Vienna: Passagen.
- Dreyfus, H. (2001). *On the internet*. New York: Routledge.
- Ess, C. (2009). Always on? Ethical and political dimensions of mobile communication technologies. In K. Nyiri (Ed.), *Engagement and exposure* (pp. 17–27). Vienna: Passagen.
- Ess, C. (2010). The embodied self in a digital age: possibilities, risks, and prospects for a pluralistic (democratic/liberal) Future? *Nordicom Information*, 32(2), 105–118.
- DET ETISKE RÅD (Danish Ethical Council) (2010). Sociale robotter-Udtalelse fra Det Etske Råd. http://etiskraad.dk/Temauniverser/Homo-Artefakt/Anbefalinger/Udtalelse%20om%20sociale%20robotter.aspx?sc_lang=en; accessed Oct. 31, 2011.
- Gallagher, S. (2003). *How the body shapes the mind*. Oxford: Oxford University Press.
- Gallagher, S. (2008). Understanding others: embodied social cognition. In P. Calvo & A. Gomila (Eds.), *Handbook of cognitive science: an embodied approach* (pp. 439–453). London: Academic Press.
- Guye-Vuilleme, A., Capin, T., Padznic, I., et al. (1999). Nonverbal communication interface for collaborative virtual environments. *Virtual Reality*, 4, 49–59.
- Heidegger, M. (1927). *Sein und Zeit*, Tübingen: Niemeyer, 15th edition 1979
- HSL (2007). *Cisco experimenting with an on-stage telepresence experience*. http://www.humanproductivitylab.com/archive_blogs/2007/11/15/cisco_experimenting_with_an_on_1.php. Accessed 24. March 2011.
- Ihde, D. (2004). A phenomenology of technics. In D. M. Kaplan (Ed.), *Readings in the philosophy of technology* (pp. 137–159). Oxford: Rowman & Littlefield.
- Jensen, L. (2003). Coming of age in a multicultural world: globalization and adolescent cultural identity formation. *Applied Developmental Science*, 7(3), 189–196.
- Kant, I. (1786). Immanuel Kant: Was heißt: sich im Denken orientieren? In: *Schriften zur Metaphysik und Logik I*. Werkausgabe hgg. von Wilhelm Weischedel. stw 188. 1. Aufl. Frankfurt/Main 1977.
- Kraut, R., Patterson, V., Lundmark, S., et al. (1998). Internet paradox: a social technology that reduces social involvement and psychological well-being? *American Psychologist*, 53(9), 1017–1031.
- Kreitler, S. (2001). An evolutionary perspective on cognitive orientation. *Evolution and Cognition*, 7, 81–97.
- Kreitler, S. (2002). Consciousness and states of consciousness. *Evolution and Cognition*, 8, 27–42.
- Kreitler, H., & Kreitler, S. (1976). *Cognitive orientation and behavior*. New York: Springer.
- Kreitler, S., & Kreitler, H. (1986a). Types of curiosity behavior and their cognitive determinants. *Archives of Psychology*, 138, 233–251.
- Kreitler, S., & Kreitler, H. (1986b). Individuality in planning: meaning patterns of planning styles. *International Journal of Psychology*, 21, 565–587.
- Kreitler, S., & Kreitler, H. (1987). Psychosemantic aspects of the self. In T. Honneth et al. (Eds.), *Self and identity: individual change and development* (pp. 338–358). London: Routledge.
- Kreitler, H., & Kreitler, S. (1990). *The cognitive foundation of personality traits*. New York: Plenum Press.
- Kreitler, H., & Kreitler, S. (2004). The motivational and cognitive determinants of defense mechanisms. *Advances in Psychology*, 136, 195–238.
- Lang, Bernhard (1963). *Martin Buber*. (Herbert Lang & Cie: Bern).
- Ling, R., & McEwen, R. (2010). Mobile communication and ethics: implications of everyday actions on social order. *Etikk i Praksis*, 4(2), 11–25.
- Mahmoud, H. (2011). Conflict defines origin: identity transformations of sudanese refugees in Cairo. *Conflict Resolution Quarterly*, 28(3), 263–289.
- Maturana, H., & Varela, F. (1987). *The tree of knowledge: the biological roots of human understanding*. Boston: New Science Library.
- Meredith, B. (2008). *Telstra chief hosts conference as hologram*. <http://www.news.com.au/telstra-chief-hosts-conference-as-hologram/story-e6frfw9-1111116449304>. Accessed 24. March 2011.
- Meteyard, L., Rodriguez Cuadrado, S., Bahrami, B., & Vigliocco, G. (in press). Coming of age: a review of embodiment and the neuroscience of semantics. DOI: [dx.doi.org/Cortex](https://doi.org/10.1016/j.cortex.2011.08.001) (Special Issue: Language and the Motor System).
- Nagy, E. (2009). The emergence of second-person cognitive science. *Trends in Cognitive Science*, 13(12), 502–504.

- Nishio, S., Ishiguro, H., & Hagita, N. (2007). Geminoid: teleoperated android of an existing person. In A. C. de Pina Filho (Ed.), *Humanoid robots: new developments* (pp. 343–352). Vienna: I-Tech Education.
- Nørskov, M. (2011). *Prolegomena to Social Robotics*, Ph.D. Dissertation at Aarhus University, Denmark.
- Noë, A. (2004). *Action in perception*. Cambridge: MIT Press.
- Ogawa, K., Nishio, S., Koda, K., Balistreri, G., Watanabe, T., & Ishiguro, H. (2011). Exploring the natural reaction of young and aged person with telenoid in a real world. *Journal of Advanced Computational Intelligence and Intelligent Informatics*, 15(5), 6.
- O’Regan, J., & Noë, A. (2001). A sensorimotor account of vision and visual consciousness. *The Behavioral and Brain Sciences*, 24, 939–1031.
- Parfit, D. (1984). *Reasons and persons*. Oxford: Oxford University Press.
- Petitot, J., Varela, F., Pachoud, B., & Roy, J.-M. (Eds.). (1999). *Naturalizing phenomenology: issues in contemporary phenomenology and cognitive science*. Stanford: Stanford University Press.
- Reddy, V. (2008). *How infants know minds*. Cambridge: Harvard University Press.
- Schilbach, L., et al. (2006). Being with virtual others: neural correlates of social interaction. *Neuropsychologia*, 44, 718–730.
- Seibt, J. (2005). Kognitive Orientierung als epistemisches Abenteuer. In W. Stegmaier (Ed.), *Orientierung—philosophische Perspektiven* (pp. 197–224). Frankfurt: Suhrkamp.
- Seibt, J. (2011). Beyond the “identity” paradigm: conflict resolution and the dynamics of self-understanding. *Conflict Resolution Quarterly*, 28(3), 229–237.
- Sicart, M. (2009). *The ethics of computer games*. Cambridge: MIT Press.
- Stutzman, F. D., & Hartzog, W. (2009). *Boundary regulation in social media* (October 8, 2009). Available at SSRN: <http://ssrn.com/abstract=1566904>.
- Tajfel, H. (1974). Social identity and intergroup behavior. *Social Science Information/sur les sciences sociales*, 13(2), 65–93.
- Tajfel, H., & Turner, J. (1979). An integrative theory of intergroup conflict. In W. Austin & S. Worchel (Eds.), *The social psychology of intergroup relations* (pp. 33–47). Monterey: Brooks/Cole.
- Turkle, S. (2011). *Alone together*. New York: Basic Books.
- Varela, F. J., Thompson, E., & Rosch, E. (1991). *The embodied mind: cognitive science and human experience*. Cambridge: MIT Press.
- Watkins, J., & Watkins, H. (1992). Ego state therapy in the treatment of dissociative disorders. In R. Kluff (Ed.), *Clinical perspectives on multiple personality disorder* (pp. 277–298). New York: American Psychiatric Pub.