

Esercizi Filosofici (2016), Vol. 11, N. 2

To what extent are content externalism and epistemic internalism compatible?

Cristina Amoretti (University of Genoa)

Inferences without necessary conclusion and easy to detect

Miguel López-Astorga (University of Talca)

Predictive engagement and motor intentionality

Valeria Bizzari (University of Pisa), Inês Hipolito (University of Wollongong)

TO WHAT EXTENT ARE CONTENT EXTERNALISM AND EPISTEMIC INTERNALISM COMPATIBLE?

Cristina Amoretti

University of Genoa

cristina.amoretti@unige.it

Abstract: The disputes between externalism and internalism in the philosophy of mind and epistemology are revealed to be greatly intertwined. In this paper, I would like to defend the compatibility between content externalism and epistemic internalism, and then explore some of its potential consequences. More precisely, I have a twofold goal: first, I shall examine the widespread idea that content externalism is not compatible with epistemic internalism, and argue that, if the two theses are appropriately understood, then there is no real tension between them; second, I shall sketch some interesting effects that may hopefully come from the combination of content externalism and epistemic internalism, as that of accommodating the constitutive link between justification and truth.

Key Words: compatibilism, content externalism, epistemic internalism, justification, privileged access.

1. *Introduction*

In the last few decades the disputes between externalism and internalism in the philosophy of mind and epistemology have revealed them to be greatly intertwined. In fact, it has been variously shown that considerations about the nature of mental content, justification and knowledge may be interdependent (Brown 2004; Goldberg 2007b, a; Ludlow and Martin 1998; Nuccetelli 2003). In what follows, I would like to defend the compatibility of content externalism and epistemic internalism, and then explore some of the possible consequences. More precisely, this paper has a twofold goal. First, I shall examine the widespread idea that content externalism is not compatible with epistemic internalism, and argue that, if the two theses are appropriately understood, then there is no real tension between them. I shall claim that if one can conceive content externalism as a diachronic holistic kind of externalism, then it happens to be compatible with epistemic internalism, no matter how one decides to typify the latter, that is, as mentalism (justification is completely determined by one's own mental states), accessibilism (justification is completely determined by elements to which one has special cognitive access), or "strong" accessibilism (in order to be justified, one must have special cognitive access to their own justificatory status). In fact, characterising content externalism as a diachronic holistic kind of externalism, one should not be worried about its

alleged incompatibility with either privileged access or introspective access to sameness and difference of content (or introspective knowledge of comparative concepts). Second, I shall sketch some interesting consequences that might hopefully come from the combination of content externalism and epistemic internalism. In this respect, I wish to argue that, by combining content externalism and epistemic internalism, one may discover a promising way to accommodate the constitutive link between justification and truth. However, I shall also point out that while this conclusion could possibly work for mentalism and accessibilism, it would be more problematic and less convincing as far as “strong” accessibilism is concerned.

2. Framing the issue

To begin with, I feel it is reasonable to embrace both content externalism and epistemic internalism. On the one hand, content externalism seems better equipped to explain, in naturalistic terms, how our mind comes to represent the external world, and how thought and language are acquired. On the other hand, epistemic internalism seems to succeed in accommodating our intuitions about the following: the kind of knowledge we wish to attribute to a cognitive mature human being; the relationship between knowledge, reason and rationality; the importance of the knower’s perspective, and the deontological character of epistemic justification. Before arguing for the compatibility of content externalism and epistemic internalism, it is crucial to clarify the theses involved in the debate.

Broadly speaking, content externalism claims that the content of (some) beliefs constitutively depends, at least in part, on objects and events in the outside world (content internalism being the thesis that the content of all beliefs completely depends on features “internal” to the subject). Of course, this is just a general definition, as many important differences should be made between: (i) physical (causal, perceptual) vs. social (linguistic); (ii) diachronic vs. synchronic; (iii) non-holistic vs. holistic; (iv) weak vs. strong, and (v) type vs. token externalism (LePore and Ludwig 2005; Amoretti 2011a). For the purpose of this paper, however, the relevant distinctions to consider are (i), (ii) and (iii), as follows.

First, physical (causal, perceptual) externalism claims that what determines the content of (some of) our beliefs is the external object or event that has actually caused it, while social (linguistic) externalism emphasises social practices, arguing that the content of (some of) our beliefs is determined by linguistic norms and conventions. Second, diachronic externalism holds that facts about our history of past causal interactions with our environment are of central importance to what content we actually have (at least to some of them). Some scholars claim that a specific past period, basically the learning period,

may be of special importance for the fixation of content, while others think that content is more or less continuously sensitive to causal interactions with the subject's environment for a long period stretching from the present into the past. Synchronic externalism argues that the content of (some of) our beliefs depends solely on our current environment and our dispositions to respond to it, so that content is not regarded as a historical fact about us, but rather as a fact about our potentialities in relation to our present surroundings. Third, non-holistic externalism maintains that the content of (some of) our beliefs depends solely on external objects and events in an atomic way, while holistic externalism holds that, even though the content of (some of) our beliefs depends on external factors, this dependence may not be its only component. Social factors and inferential relations among beliefs may be also fundamental in determining the content of a particular belief. This two-fold dependency implies that a difference in external factors must lead to a change of content, while the sameness of external factors does not automatically lead to the identity of content.

In what follows I will try to show that diachronic holistic content externalism (DHCE) has all the resources to be compatible with epistemic internalism.

(DHCE) Diachronic Holistic Content Externalism: the content of (some) beliefs constitutively depends on the causal history that has tied up the subject with a certain external object or event in connection with various other beliefs.

In other words, DHCE holds that mental content constitutively depends on the set of those specific circumstances in which the subject has learned and effectively used the beliefs at issue, circumstances which include external objects and events, social interactions, and relations with various other beliefs.

Moving to epistemology and epistemic internalism, a preliminary distinction between internalism about knowledge and justification should be made. The former simply claims that knowledge requires justification, while the latter holds that justification should be understood "internally" (externalism about knowledge being the thesis that knowledge does not require justification, and externalism about justification that justification should be understood "externally"). Here I shall refer to internalism about justification only, so epistemic internalism would be the thesis that justification is completely determined by "internal" conditions.

Another important difference concerns the kind of justification one wishes to consider. On the one hand, propositional justification is the justification for believing that p (which is like having a good reason for believing that p). On the other hand, doxastic justification means being justified in believing that p (or having a justified belief, or basing one's belief on reasons). As doxastic justification may clearly depend on facts external to the subject (as in the causal

relations between beliefs and reasons), many scholars have pointed out that epistemic internalism should merely focus on propositional justification (Greco 2005; Fumerton 2007). I stick with that conclusion, so I shall refer to propositional justification only.

Given the above distinctions, epistemic internalism is the claim that the justification for a belief that *p* is completely determined by some features that are “internal” to the subject’s mind. To put it another way, whether a subject is justified in believing that *p* supervenes on some elements internal to the subject’s mind. This thesis, however, can be refined at least into three different claims, which have been dubbed *mentalism* (M), *accessibilism* (A), and “*strong*” *accessibilism* (SA).

(M) **Mentalism:** justification is completely determined by one’s own mental states.

M is a rather weak claim, as it simply holds that one’s own mental states (beliefs, propositional attitudes, perceptual states, and so on) completely determine the justificatory status of one’s particular belief that *p* (Conee and Feldman 2001). This means that even non-reflectively accessible mental states – as far as they are considered mental – can determine the justificatory status of one’s particular belief.

(A) **Accessibilism:** justification is completely determined by one’s own reflectively accessible states (that is by some elements to which the subject has special cognitive access).

To put it differently, A argues that reflectively accessible states completely determine the justificatory status of one’s particular belief that *p*. Talking about “accessibility” and “special access”, scholars typically refer to what is available from introspection or reflection alone, or *a priori*, without inference from observation of one’s behaviour, speech or environment (Audi 2010). Given this definition of A, the reason why focusing on propositional justification instead of doxastic justification becomes clear: while the causal origin of a belief could be inaccessible from introspection or reflection alone, the justificatory relationships between beliefs (such as whether or not *p* contradicts *q*) are instead accessible.

(SA) “**Strong**” **Accessibilism:** one’s own justificatory status must be reflectively accessible as well.

This latter thesis is the strongest claiming that, in order to have a justification for believing that *p*, the subject must have special access to her own justificatory status. This implies that, if the subject is unable to reflectively determine whether or not her belief that *p* is justified, then she is not justified in believing

that *p*. Again, the notion of special access refers to what is accessible by introspection or reflection alone, but it is important to note that the notion of special access should not be considered as referring to either direct or infallible knowledge, as coming to know one's own justificatory status certainly requires some reasoning. Even if it is quite demanding, SA probably best exemplifies our deep intuitions about epistemic internalism.

In the next section, I shall argue that, if we characterise content externalism as DHCE, then it would be compatible with epistemic internalism, no matter how the latter is typified, that is, as M, A, or SA.

3. Some anti-compatibilist arguments

Starting with M, it is easy to see that anti-compatibilism cannot get off the ground. If the basic tenet of epistemic internalism is that justification is completely determined by one's own mental states, then it is clearly compatible with DHCE (and, more generally, with content externalism), as beliefs are definitely mental states, however their content would actually be determined. As Earl Conee points out, if epistemic internalism "is the thesis that for epistemic purposes the 'internal' is the mental", then "since content externalism expands the factors that fix the mental, content externalism expands the supervenience base for justification according to mentalism" (Conee 2007: 51). We can therefore easily state that DHCE is compatible with M.

The alleged incompatibility between content externalism and A, and/or SA, depends on the widespread opinion that content externalism undermines the thesis of Privileged Access (PA), that is, the subject's ability to access the content of her own beliefs by introspection or reflection alone, or a *priori*, without inference from observation of her own behaviour, speech or environment. To put it generally, content externalism claims that (some) mental contents are determined, at least in part, by external factors, but those factors, being external, may be inaccessible to the subject from introspection or reflection alone; (some of) the subject's mental contents, then, may as well be inaccessible from introspection or reflection alone. If the subject may be unable to access the content of (some of) her own beliefs by introspection or reflection alone, then PA is denied. As a consequence, the subject may also ignore the content of some of her own beliefs, and thus be mistaken about what she actually believes.

The point is that the lack of PA undercuts both A and SA. The two arguments for incompatibilism run as follows. Against A: given content externalism, the subject may lack the ability to access the content of (some of) her own beliefs by introspection or reflection alone; thus, the subject may lack the ability to access by introspection or reflection alone (some of) the factors determining the justificatory status of (some of) her beliefs. Hence, some

introspectively or reflectively inaccessible states do contribute to determining the justificatory status of (some of) her beliefs. As justification is not completely determined by the subject's own accessible states, then A is false. Against SA: given content externalism, the subject may lack the ability to access by introspection or reflection alone the content of (some of) her own beliefs; thus, the subject may lack the ability to access by introspection or reflection alone her own justificatory status. As the subject's own justificatory status may be introspectively or reflectively inaccessible, then SA is false.

Generally speaking, it is not immediately clear whether content externalism undermines PA. Externalists generally claim that to rebut this objection it would be sufficient to note that what determines the content of our first-order beliefs (I believe that p), whatever it is, is also what determines the content of our corresponding second order beliefs (I believe to believe that p), leaving no room for error (Burge 1988, 1996; Heil 1988; Davidson 1987; Sainsbury and Tye 2012). Then, although the subject may obviously have false beliefs, she would still retain PA to the content of her own beliefs.

Two arguments have been introduced to resist this compatibilist solution. First, let us consider the likelihood of slow switching, of which the subject is not aware, between Earth and Twin Earth (Boghossian 1989; Goldberg 1997, 1999; Ludlow 1995, 1997). In this case, the possibility that a sample of transparent and odourless liquid is twin-water, instead of water, becomes a relevant alternative in order to evaluate whether a subject has PA to her own water-beliefs. Given the likelihood of slow switching and the fact that they are unnoticed by the subject, she would not be able to establish by introspection or reflection alone whether she has water-beliefs or twin-water-beliefs. This means that the subject would not be able to determine by introspection or reflection alone what her own current beliefs are. The argument runs as follows:

- (P1) Before the switch, on Earth, the subject has water-beliefs;
- (P2) After the switch, on Twin-Earth, the subject has twin-water-beliefs;
- (P3) The subject is unaware of the switch and cannot distinguish Earth from Twin-Earth by introspection or reflection alone;
- (P4) The subject does not know (or is unable to determine) by introspection or reflection alone whether she has water-beliefs or twin-water-beliefs;
- (C) Thus, the subject cannot reflectively access the content of her own beliefs.

Such a conclusion obviously implies that the subject would not be able to introspectively or reflectively access (some of) the states that determine the justificatory status of (some of) her own beliefs. Thus, some introspectively or reflectively inaccessible states would possibly determine the justificatory status of (some of) the subject's beliefs, which means the denial of A. Also, the subject

would not be able to introspectively or reflectively access her own justificatory status too, which means the denial of SA.

Against this argument some scholars have pointed out that, even if the subject is unable to determine by introspection or reflection alone whether she has water-beliefs or twin-water-beliefs, this does not undermine PA (Falvey and Owens 1994). Moreover, it has been noted that the current belief that p counts as evidence for the belief of believing that p , and this would be enough to rule out all possible alternatives, relevant or not (Sainsbury and Tye 2012).

Finally, the success of the argument strongly depends on what kind of content externalism is defended. According to DHCE, what determines content is the subject's causal history (together with holistic constraints). This means that, immediately after the switch, the subject would still have water-beliefs (at least for a certain amount of time), and the second premise would simply be false. Of course, as the subject stays on Twin-Earth long enough, the contents of her beliefs would change in relation to the new environment. The point is how exactly these contents would change (Brown 2004). Some scholars argue that, after the switch, the subject would come to have both water-beliefs and twin-water-beliefs, and thus defend the "two-concepts view" (Burge 1998; Gibbons 1996). Others think that, after the switch, the content would change from water to twin-water, and thus maintain the "replacement view" (Bernecker 1998; Brueckner 1997). However, as I argued elsewhere (Amoretti 2007, 2011b), there is a third alternative that can be dubbed "amalgam view" and better fits with DHCE: after the switch the content of the subject's belief would slowly change, as new causal interactions with the novel environment take place, having a sort of "amalgam" concept referring to both water and twin-water. In this case, the subject would still be able to introspectively or reflectively access not only the states determining the justificatory status of her own beliefs, but also their own justificatory status, which means that DHCE would still be compatible with both A and SA.

Another anti-compatibilist argument is based on a *reductio* (McKinsey 1991). Let us assume that the subject knows (or is able to determine) by introspection or reflection alone the following premises:

- (P1) She believes that water is wet [given PA];
- (P2) If she believes that water is wet, then some external fact holds [given content externalism];
- (C) Hence, some external fact holds.

This would mean that the subject knows (or is able to determine) by introspection or reflection alone that some external fact holds, such as that water exists; but this conclusion would be absurd. As a consequence, holding content externalism would lead to abandon PA, and then to admit the incompatibility with both A and SA.

This argument can be resisted in various ways. First, many content externalists would simply deny that (P2) can be known by introspection or reflection alone. Moreover, even admitting that (P2) can be known by introspection or reflection alone, the *reductio* can be resisted, as additional premises would be needed. According to Colin McGinn (1989), it would be necessary to add a premise stating that the subject also believes that the belief that water is wet contains an atomic, natural kind concept, but this can hardly be known *a priori*. Tyler Burge (1979) similarly argues that it would be necessary to add a premise asserting that the subject also believes that the belief that water is wet contains a deferential concept referring to a certain linguistic community; again, this cannot be known *a priori*. Finally, Donald Davidson (1990, 1991, 1995, 1997) would probably maintain the necessity of adding a premise claiming that the subject also believes that the belief that water is wet contains a perceptual concept that “anchors” mind and language to the world, and this claim too can hardly be known *a priori* (see Amoretti 2008; 2011). So, to conclude the discussion on the compatibility of PA with content externalism, we can state that DHCE is compatible with both A and SA.

Another anti-compatibilist strategy is to argue merely against the compatibility between content externalism and SA. This move is based on the conviction that content externalism undermines the thesis of Introspective Knowledge of Comparative Concepts (IKCC) – the basic issue clearly being that the lack of IKCC would undermine SA. IKCC can be seen as the conjunction of two theses:

- i) Transparency of sameness of content (TSC): for any couple of a subject’s beliefs that *p* and that *q* entertained at a certain time *t*, if the beliefs that *p* and that *q* have the same content, then at *t* the subject knows by introspection or reflection alone that they have the same content.
- ii) Transparency of difference of content (TDC): for any couple of a subject’s beliefs that *p* and that *q* entertained at a certain time *t*, if the beliefs that *p* and that *q* have a different content, then at *t* the subject knows by introspection or reflection alone that they have a different content.

Given content externalism, a subject with two occurring beliefs, that *p* and that *q*, may not be able to establish by introspection or reflection alone whether the beliefs that *p* and that *q* have the same (or a different) content. If this is the case, then the subject cannot introspectively or reflectively grasp all the logical relations between these beliefs. As logical relations affect the justificatory status of one’s beliefs, the subject would not introspectively or reflectively know whether (some of) her beliefs are justified, and thus she would lack *a priori* access to the justificatory status of (some of) her beliefs. Accordingly, SA would be false (even if the subject might not be able to establish by introspection or reflection alone whether *p* and *q* have the same or a different

content, this does not affect A, as long as one admits that the subject has access to her own contents by introspection or reflection alone).

Two arguments have been developed to support the anti-compatibilist conclusion. First, let us suppose that, even if cilantro and coriander are clearly the same herb, the subject does not know that. The argument would run as follows:

- (P1) The subject believes that cilantro is used in Mexican dishes [hypothesis];
- (P2) The subject believes that coriander is used in Mexican dishes [hypothesis];
- (P3) The two beliefs have the same content [given content externalism];
- (P4) As the subject ignores that cilantro and coriander are the same herb, she cannot establish by introspection or reflection alone that the two beliefs have the same content [as the subject needs to make an empirical investigation to discover the identity between cilantro and coriander];
- (C) Thus, the subject cannot introspectively or reflectively grasp all the logical relations between her beliefs.

This conclusion would imply that the subject may lack introspective or reflective access to the justificatory status of (some of) her beliefs. Accordingly, SA would be false. There are various ways to resist this reasoning, but endorsing DHCE allows one to reject (P3), that is the idea that the belief that cilantro is used in Mexican dishes and the belief that coriander is used in Mexican dishes have the very same content. According to DHCE, what gives content to our beliefs is not the external cause by itself, but rather the specific causal history that has tied up the subject with a certain external object in connection with many other beliefs. To put it differently, the content is determined by the set of those specific circumstances, encompassing both causal chains and holistic relationships, in which the subject has acquired the concepts and effectively learned the beliefs at issue. As I argued elsewhere, if this externalist position holds, then there would be two different beliefs and thus no sameness of content where the anti-compatibilists see it (Amoretti 2007).

The second argument, which is based on the likelihood of slow switches and the fact that they are unnoticed by the subject, runs as follows:

- (P1) At *t*, the subject believes that, as a child, she lived near a river full of spring water [she were on Earth and thus, given externalism, her belief refers to water];
- (P2) At *t*, the subject believes that now she lives near a river full of spring water [she is now on Twin-Earth and thus, given externalism, her belief refers to twin-water];
- (P3) The two beliefs have a different content [given externalism];

(P4) As the subject is unaware of the switch and cannot distinguish Earth from Twin-Earth, she cannot establish by introspection or reflection alone that the two beliefs have a different content [as the subject needs to make an empirical investigation to distinguish Earth from Twin-Earth];

(C) Thus, the subject cannot introspectively or reflectively grasp all the logical relations between her beliefs.

This conclusion would obviously imply that the subject may lack introspective or reflective access to the justificatory status of (some of) her beliefs. Accordingly, SA would be false. In this respect, we have already seen that the plausibility of this line of argument strongly depends on how the subject's contents would actually change: only the "two-concepts view" would be challenged by the objection. On the contrary, holding the "amalgam view", the option that better fits with DHCE, the objection is immediately blocked. This means that DHCE is still compatible with SA.

More generally, and more importantly, there are some strong doubts about the very plausibility of IKCC. If this is so, then the possibility that content externalism would undermine IKCC would not be a problem for compatibilism itself. According to some scholars, requiring IKCC would be too strong a request, as there are some important analogies with perception (Sainsbury and Tye 2012). Let us think about illusions (as the Muller-Lyer illusion): in similar cases, we could perceive the same (or a different) object without recognising that it is the same (or a different) object. Moreover, the notion of rationality that is implicit in IKCC seems to be too strong: as sameness and difference of content can be determined through empirical investigation, it does not seem legitimate to pretend that they must be judged by introspection or reflection alone. The very philosophers who criticise compatibilism agree that rationality does not depend on the number of true beliefs that a subject holds, but on the logical relations linking these beliefs. And logical relations, by definition, leave aside the content of the beliefs involved. For instance, a subject would be irrational if, believing that $a \rightarrow b$ and $b \rightarrow c$, she fails to recognise that $a \rightarrow c$; but she should be regarded as rational if she believes that $a \rightarrow b$ and $c \rightarrow d$ but, ignoring that $b = c$, fails to recognise that $a \rightarrow d$, as the fact that $b = c$ (or $b \neq c$) can only be determined through empirical investigation, not by introspection or reflection alone.

For example, let us assume that a subject believes both that someone shot John Locke, and that nobody shot John Locke. It would be essential to distinguish cases where the subject is aware to hold contradictory beliefs (in this case she should be judged irrational) from cases where the subject ignores to hold contradictory beliefs, for instance, believing that someone shot John Locke (taking this claim to refer to the character of Lost), and that nobody shot John Locke (mistakenly taking this claim to refer to the philosopher). In this latter case the subject would believe falsely, but rationally and would be justified in

doing so, that someone shot John Locke and that nobody shot John Locke (Sainsbury and Tye 2012). When the aim is judging the subject's own rationality, it does not matter if she has true or false beliefs. The subject can legitimately make some inferences that are coherent with her overall knowledge, even if they are wrong from a metaphysical perspective. These inferences, in fact, can be perfectly compatible with the subject's whole set of beliefs, even though they lead to a basically wrong conclusion due to the fact that our perspective is limited and fallible. If we take as a criterion for rationality the transparency of mental content, we fail to consider one fundamental characteristic of doxastic contexts, i.e., the fact that they are opaque.

4. Some consequences of compatibilism

Having granted compatibilism between content externalism and epistemic internalism (in its various forms), it is now possible to evaluate some of its consequences. In this respect, I wish to argue that, by combining content externalism and epistemic internalism, one may discover an interesting way to accommodate the constitutive link between justification and truth. However, I shall point out that this conclusion may work for M and A, but it is more problematic and less convincing as far as SA is concerned.

According to some critics, epistemic internalism would fail to accommodate the link between justification and truth, as it does not explain how justification relates to truth, or even denies the necessity of any relation between justification and truth; whether one's belief is likely to be true is not determined by either one's own mental states or one's own introspectively or reflectively accessible states. However, granting the link between justification and truth would be important for epistemic internalism, as this would help to distinguish epistemic justification from non-epistemic justification (such as prudential, pragmatic or moral justification). Content externalism may help to accommodate this missing link by showing that beliefs are, by their own nature, veridical (Davidson 1990).

Broadly speaking, some scholars (for a review, see Brueckner 2016) have variously argued that content externalism, possibly combined with PA, may have some important epistemological consequences, such as:

- i) that we are not brains in a vat;
- ii) that a coherent set of beliefs cannot be totally or even largely false;
- iii) that it is highly probable that a coherent set of beliefs is not totally or even largely false;
- iv) that perceptual beliefs cannot be totally or even largely false, and
- v) that it is highly probable that perceptual beliefs are not totally or even largely false.

Let us consider DHCE: the content of (some) beliefs is determined by the causal history that has tied up the subject with a certain external object or event in connection with various other beliefs. This seems to imply that a coherent set of beliefs cannot be totally or even largely false. To put it another way, if one has beliefs at all, then DHCE guarantees that a set of beliefs cannot be totally or largely false, in particular as far as perceptual beliefs are concerned. Probably, as I tried to show in previous works (Amoretti 2011b, 2008b, a), DHCE has weaker consequences, such as:

- i) that a coherent set of beliefs cannot always have been totally or even largely false, or
- ii) that it is highly probable that a coherent set of beliefs would not always have been totally or even largely false.

To begin, let us apply the above consequences of DHCE to M (justification is completely determined by one's own mental states) and A (justification is completely determined by one's own reflectively accessible states). With some simplifications, we can say that in both cases justification is completely determined by reasons in the form of other (reflectively accessible) beliefs. Two classic theories of justification are typically regarded as internalist in this sense, that is, coherentism and foundationalism.

According to coherentism, all justified beliefs are inferentially justified, that is they cohere with a significant number of other beliefs. The basic idea is that there is a presumption in favour of the truth of a belief that coheres with a vast array of other beliefs, and every belief thus counts as justified in the light of such a presumption. However, a set of (at least) *largely true* beliefs is needed to secure the above presumption, and coherence alone cannot guarantee that a coherent set of beliefs is largely true (Davidson 1983, 1990). On the contrary, content externalism seems to be able to do the job granting at least that a coherent set of beliefs cannot be (or cannot always have been) totally or even largely false – or that it is highly probable that a coherent set of beliefs is not (has not always have been) totally or even largely false. If a belief coheres with a vast array of beliefs, which is not totally or even largely false, then the presumption in favour of its truth, and thus its justification, is secured.

A similar reasoning can be made for foundationalism, according to which all justified beliefs ultimately rest upon a foundation of non-inferentially justified beliefs, such as perceptual beliefs (or self-evident beliefs, etc.). Here the basic idea is that there is a presumption in favour of the truth of a belief that rest upon some non-inferentially justified beliefs, such as perceptual beliefs (which derive their justification from some non-doxastic perceptual states), and every belief thus counts as justified in the light of such a presumption. However, perceptual beliefs must be largely true to secure the above presumption, and perception alone cannot guarantee that. Again, content externalism seems at least to be able

to grant that one's own perceptual beliefs cannot be totally or even largely false – or that it is highly probable that one's own perceptual beliefs are not totally or even largely false. If a belief rests upon non-inferentially justified beliefs, such as perceptual beliefs, which are not totally or even largely false, then the presumption in favour of its truth, and thus its justification, is secured.

In both cases, content externalism seems to be able to provide the link between justification and truth that was missing from internalist theories of justification, at least conceiving them as M or A. But what about SA? I feel that, in this case, the above results would be useless. According to SA, a subject must be aware by introspection or reflection alone of her justificatory status. This means that she must be aware, by introspection or reflection alone, not only that a belief that *p* coheres with many other coherent beliefs, but also that there are good reasons to assume that cohering beliefs cannot be totally or even largely false (alternatively, not only that a belief that *p* rests upon some non-inferentially justified beliefs such perceptual beliefs, but also that there are good reasons to assume that perceptual beliefs cannot be totally or even largely false). This would be quite difficult, as it requires that the subject knows content externalism and its consequences, which is highly improbable.

To conclude, at least some kinds of content externalism, such as DHCE, are compatible with epistemic internalism in all its forms. By combining content externalism and epistemic internalism, one may discover an interesting way to accommodate the constitutive link between justification and truth – a way that can hardly be provided by epistemic internalism alone. However, this result obtains for M and A, while it is more problematic and unconvincing as far as SA is concerned.

Acknowledgements. I gratefully thank Nicla Vassallo for her helpful and valuable comments on an earlier draft of this paper.

References

- AMORETTI, M.C.
2007 «Triangulation and Rationality», *Epistemologia* 2, 307-326.
2008a «Davidson, Self-Knowledge, and Scepticism», in M.C. Amoretti and N. Vassallo (eds.), *Knowledge, Language, and Interpretation: On the Philosophy of Donald Davidson*, Ontos Verlag, Frankfurt.
2008b *Il triangolo dell'interpretazione: sull'epistemologia di Donald Davidson*, FrancoAngeli, Milano.
2011a *La mente fuori dal corpo. Prospettive esternaliste in relazione al mentale*, FrancoAngeli, Milano.
2011b «Triangulation between Externalism and Internalism», in M.C. Amoretti and G. Preyer (eds.), *Triangulation: From an Epistemological Point of View*, Ontos Verlag, Frankfurt, 47-68. Frankfurt.

AUDI, R.

2010 *Epistemology: A Contemporary Introduction to the Theory of Knowledge*, 3rd ed., Routledge, London.

BERNECKER, S.

1998 «Self-Knowledge and Closure», in P. Ludlow and N. Martin (eds.), *Externalism and Self-Knowledge*, CSLI Publications, Stanford, CA, 333-350.

BOGHOSSIAN, P.

1989 «Content and Self-Knowledge», *Philosophical Topics* 17, 5-26.

BROWN, J.

2004 *Anti-individualism and Knowledge*, MIT Press, Cambridge, MA.

BRUECKNER, A.

1997 «Externalism and memory», *Pacific Philosophical Quarterly* 78, 1-12.

2016 «Skepticism and Content Externalism» (previously entitled «Brains in a Vat»), in E.N. Zalta (ed.), *The Stanford Encyclopedia of Philosophy* (Winter 2016 Edition), URL = <<https://plato.stanford.edu/archives/win2016/entries/skepticism-content-externalism/>>.

BURGE, T.

1979 «Individualism and the Mental», *Midwest Studies in Philosophy* 4, 73-122.

1988 «Individualism and Self-Knowledge», *The Journal of philosophy* 85, 649-663.

1996 «Our Entitlement to Self-Knowledge», *Proceedings of the Aristotelian Society* 96, 91-116.

1998 «Memory and Self-Knowledge», in P. Ludlow and N. Martin (eds.), *Externalism and Self-Knowledge*, CSLI Publications, Stanford, CA, 351-369.

CONEE, E.

2007 «Externally Enhanced Internalism», in S. Goldberg (ed.), *Internalism and Externalism in Semantics and Epistemology*, Oxford University Press, Oxford, 51-67.

CONEE, E. AND FELDMAN, R.

2001 «Internalism defended», in H. Kornblith (ed.), *Epistemology: Internalism and Externalism*, Blackwell, Oxford, 231-259.

DAVIDSON, D.

1983 «A Coherence Theory of Truth and Knowledge», in D. Henrich (ed.), *Kant oder Hegel*, Klett-Cotta, Stuttgart, 423-438.

1987 «Knowing One's Own Mind», *Proceedings and Addresses of the American Philosophical Association* 60, 441-458.

1990 «Epistemology Externalised», in D. Davidson, *Subjective, Intersubjective, Objective*, Oxford University Press, Oxford, 193-203.

1991 «Three Varieties of Knowledge», *Philosophy* 66, 156-166.

1995 «The Problem of Objectivity», *Tijdschrift voor Filosofie* 57, 203-220.

1997 «The Emergence of Thought», in D. Davidson, *Subjective, Intersubjective, Objective*, Oxford University Press, Oxford, 123-134

FALVEY, K. AND OWENS, J.

1994 «Externalism, Self-knowledge, and Skepticism», *Philosophical Review* 103, 107-137.

GIBBONS, J.

1996 «Externalism and Knowledge of Content», *Philosophical Review* 105, 287-310.

GOLDBERG, S.

1997 «Self-Ascription, Self-Knowledge, and the Memory Argument», *Analysis* 57, 211-219.

1999 «Word-Ambiguity, World-Switching, and Knowledge of Content: Reply to Brueckner», *Analysis* 59, 212-217.

2007a *Anti-Individualism. Mind and Language, Knowledge and Justification*, Cambridge University Press, Cambridge, MA.

2007 *Internalism and Externalism in Semantics and Epistemology*, Oxford University Press, Oxford.

HEIL, J.

1988 «Privileged Access», *Mind* 97, 238-251.

LEPORE, E., AND LUDWIG, K.

2005 *Donald Davidson. Meaning, Truth, Language, and Reality*, Clarendon Press, Oxford.

LUDLOW, P.

1995 «Externalism, Self-Knowledge, and the Prevalence of Slow-Switching» *Analysis* 55: 45-49.

1997 «On the Relevance of Slow Switching» *Analysis* 57: 285-286.

LUDLOW, P., AND MARTIN, N. (EDS.)

1998 *Externalism and Self-Knowledge*, CSLI Publications, Stanford, CA.

MCGINN, C.

1989 *Mental Content*, Blackwell, New York, NY.

NUCCETELLI, S.

2003 *New Essays on Semantic Externalism and Self-Knowledge*, MIT Press, Cambridge, MA.

SAINSBURY, R.M. AND TYE, M.

2012 *Seven Puzzles of Thought and How to Solve Them: An Originalist Theory of Concepts*, Oxford University Press, Oxford.

INFERENCES WITHOUT NECESSARY CONCLUSION AND EASY TO DETECT

Miguel López-Astorga

Institute of Humanistic Studies “Juan Ignacio Molina”

University of Talca

milopez@utalca.cl

Abstract: Aristotle indicates four types of inferences whose conclusion does not necessarily follows its premises. In this paper, I review them from the mental models theory and try to show that, if this last framework is correct, people should note without a lot of effort that their conclusions are not necessary. I mainly base my analysis on the distinction between canonical and noncanonical models provided by Khemlani, Lotstein, Trafton, and Johnson-Laird (2015).

Key Words: Aristotelian logic, inference, mental models, necessary conclusion, quantification.

1. *Introduction*

The mental models theory (e.g., Johnson-Laird 2010; 2012; 2015; Johnson-Laird, Khemlani, & Goodwin 2015; Khemlani, Lotstein, Trafton, & Johnson-Laird 2015; Khemlani, Orenes, & Johnson-Laird 2012; 2014; Orenes & Johnson-Laird 2012; Quelhas, Johnson-Laird, & Juhos 2010; Ragni, Sonntag, & Johnson-Laird 2016) is a contemporary theory on reasoning essentially based on the idea that human beings make inferences considering possibilities sets related to sentences in natural language. Several aspects of Aristotelian logic have been analyzed from the theoretical machinery of this approach (e.g., López-Astorga 2016a; 2016b; 2016c), and the results refer to different ideas. One of them that is very relevant is that, if it is true that the mental models theory (from now on, MMT) really describes and explains human reasoning, certain particular aspects of Aristotle’s logical system can be used as instruments to predict the conclusions that will be deduced in some types of inferences, as well as the inferential mental activities that will be easier or more difficult for people. This is so for two reasons. Firstly, MMT seems to be able to show to what extent the structures in Aristotelian logic require cognitive effort to be understood or applied. Secondly, it appears that certain correspondences between the syllogisms considered correct by Aristotle and the results expected by MMT in those very syllogisms exist.

This paper is intended to continue in this way and to review four inferences that, according to Aristotle, do not have a necessary conclusion. Thus, its aim is to check whether, following MMT, it must be expected that individuals quickly

note that the conclusions of such inferences are not guaranteed by their premises or, on the contrary, that it is very hard to be aware of that fact.

So, this paper can be considered to be a complement of different works with similar goals, but especially of those akin to that of López-Astorga (2016a), in which it is argued that MMT proves that all the moods of the four figures in Aristotelian logic represent schemata that people should apply without difficulties. The challenge here is, nevertheless, to analyze whether or not the situation is as clear in the case of certain incorrect syllogisms as in that of, for example, the moods and figures. To do that, I will begin by commenting on the particular theses of MMT that need to be taken into account in a study of this kind. Then I will describe the exact four syllogisms to which I am referring. Finally, I will indicate what MMT can say with regard to those four inferences.

2. *MMT and quantification*

As pointed out, the literature on MMT is not scarce. However, a very relevant paper about the way it addresses the quantified sentences (i.e., the type of sentences that is used in Aristotelian logic) is evidently that of Khemlani *et al.* (2015). I will primarily base this section on that paper and on the explanation of it given by López-Astorga (2016a: 6ff).

MMT claims that what is important in reasoning and the inferential activity is not the logical form of sentences, but the possibilities that such sentences enable (e.g., Johnson-Laird 2010). In this way, according to the theory, people think about the possibilities of each sentence by means of representations that are iconic in a sense akin to that given by Peirce (1931/1958) to that word (see, e.g., Johnson-Laird 2012). So, following general assumptions such as these ones, Khemlani *et al.* state that, when faced with quantified sentences, people obviously consider the combinations of possibilities that can be assigned to them too. The only problem in this regard is that, as also indicated by the theory for sentences in general, effort is needed to identify all of the possibilities and, for this reason, individuals sometimes only pay attention to the ones that are easier to detect. Thus, the particular expressions used by them in their paper are “canonical models” to refer to the sets of possibilities not requiring a lot of effort to be captured, and “noncanonical models” to mention the other ones. In their Table 1 (Khemlani *et al.* 2015: 2077), they show examples of canonical and noncanonical models for different kinds of quantified sentences. Nevertheless, for my aim here, it is only necessary to mainly take the canonical models of the affirmative and negative universal sentences into account. Starting with the affirmative universal ones, that is, those that can be formally expressed as [A(sp)] (where, obviously, “A” indicates that the sentence is affirmative universal, “s” means “subject”, and “p” represents “predicate”), which claim

that “all of the *s* are *p*”, it can be said that Khemlani *et al.* (2015) give an example of canonical model similar to the following:

s & *p*
s & *p*
s & *p*

Each of these three possibilities describes an element. In this case, the three elements have the same characteristics (all of them have both the property “*s*” and the property “*p*”), which is consistent with what a formula such as [*A*(*sp*)] means. Of course, further effort can be made and people can come to other noncanonical models by noting, for example, that an element without the property “*s*” but with the property “*p*” ($\neg s \ \& \ p$, where “ \neg ” is negation) or an element without the property “*s*” and without the property “*p*” ($\neg s \ \& \ \neg p$) are also compatible with the asseveration that “all of the *s* are *p*”. Nonetheless, according to MMT, only a model such as that indicated, i.e., a model in which all the elements have both properties, can be quickly and easily detected given an affirmative universal sentence.

As said, the other type of quantified sentence that is interesting in this paper is the one of the negative universal sentences, that is, of those that can be formally expressed as [*E*(*sp*)] (where “*E*” means that the sentence is negative universal), which claim that “none of the *s* is *p*”. In this case, the Khemlani *et al.*’s example is akin to this one:

s & $\neg p$
s & $\neg p$
 $\neg s \ \& \ p$

Now, the first and the second elements have only the property “*s*”, and not the property “*p*”. On the other hand, the third element does not have the property “*s*”, but it does have the property “*p*”. It is obvious that cognitive effort can reveal other possibilities here as well. However, as mentioned, the canonical models are the most relevant for my goals. This is so because, as also indicated, what I wish to check is whether or not people can note without effort that certain inferences rejected by Aristotle, indeed, are not correct. In this way, if just the canonical models described are enough to show that such inferences are not valid, it will be possible to state that it is not difficult for people to be aware of this last fact, at least if MMT is assumed. Furthermore, as also pointed out, the models that Khemlani *et al.* (2015) assign in Table 1 to other types of quantified inferences are not important here either, since, as accounted for below, the four inferences that will be reviewed have the same two premises, one of them being

an affirmative universal sentence and the other one being a negative universal one.

3. Four inferences with conclusions that are not necessary

The four inferences are exactly in *Ἀναλυτικὰ Πρότερα* (*Analytica Priora*) 26a2-9, and have been studied in different works too (e.g., Johnson 2004; Lukasiewicz 1957). Following the Lukasiewicz's (1957: 21) idea that Aristotelian logic was 'implicational', Johnson (2004: 255) proposes this general structure for them:

$$A(mp) \rightarrow [E(sm) \rightarrow x]$$

Where, evidently, “ \rightarrow ” denotes implication and “m” refers to the middle term.

Thus, the difference between them is the content of [x], which can be one of these formulae:

I(sp) (where “I” means that the sentence is affirmative particular, i.e., that it expresses that “some of the s are p”)

O(sp) (where “O” indicates that the sentence is negative particular, i.e., that it expresses that “some of the s are not p”)

A(sp)

E(sp)

So, the four inferences are formally as follows:

$$A(mp) \rightarrow [E(sm) \rightarrow I(sp)]$$

$$A(mp) \rightarrow [E(sm) \rightarrow O(sp)]$$

$$A(mp) \rightarrow [E(sm) \rightarrow A(sp)]$$

$$A(mp) \rightarrow [E(sm) \rightarrow E(sp)]$$

As said, according to Aristotle, the conclusions, that is, the content of [x] or the consequents of the consequents in the previous formulae, do not necessarily follow from the premises (i.e., from [A(mp)] and [E(sm)]) in these four cases. As explained by Johnson (2004), they are also rejected in the system proposed by Lukasiewicz (1957). Nevertheless, what is interesting for this paper is

whether or not, following MMT, it is easy for human beings without logical background to see that. Because we know which the canonical models of sentences such as [A(sp)] and [E(sp)] are in that theory, it should not be difficult to clarify this point. I try to do it in the next section.

4. *Would it be hard for people to reject the inferences?*

As I will show now, from the MMT perspective, the answer to this question is clearly negative. And this is so because, to understand that the conclusions do not follow from the common premises, it is only necessary to take the canonical models of such premises into account. Given that the first one is [A(mp)], its canonical model can be as follows:

[I]: $m \ \& \ p$

[II]: $m \ \& \ p$

[III]: $m \ \& \ p$

Likewise, the one of the second premise can be this:

[IV]: $s \ \& \ \neg m$

[V]: $s \ \& \ \neg m$

[VI]: $\neg s \ \& \ m$

Hence we have six possibilities, [I], [II], [III], [IV], [V], [VI]. Accordingly, what must be done then is to check whether or not what is expressed by the four conclusions matches the description of the domain or universe given by such possibilities. In this way, it can be said that the first conclusion [I(sp)] would follow the premises if at least one element with both the property “s” and the property “p” were found. But we cannot do that because in [I], [II], and [III], while the elements have the property “p”, it is not known whether or not they have the property “s” as well. There is no information about that and, even though it is possible that all of them have the property “s”, it is also possible that none of them has it. Something similar can be said about [IV] and [V]. They represent elements with “s”, but we do not know whether or not they have “p” too. Furthermore, in [VI] the element does not have the property “s” and, therefore, it is not useful to confirm the conclusion. Thus, given the premises, [I(sp)] is not necessarily true.

The situation is even clearer in the case of the second conclusion. Indeed, [O(sp)] would be a valid conclusion if there were at least a case of “s” without “p”. Nevertheless, again, it is not possible to know for sure whether or not there is such a case. [IV] and [V] are the only elements that, undoubtedly, have the property “s”, but it is unknown whether or not they have the property “p”.

Maybe they do not have it but it is also possible that they do. On the other hand, [I], [II], and [III] have “p”, and [VI] does not have “s”.

As far as the third conclusion [A(sp)] is concerned, it is necessary that all the elements with “s” have “p” as well. In [I], [II], and [III], we do not know whether or not they have “s”. In [IV] and [V], what is unknown is whether or not they have “p”. And, in [VI], it is absolutely clear that “s” is there.

Finally, the case of the fourth conclusion is obvious too. We could accept [E(sp)] if all of the “s” were “¬p”. Nonetheless, [I], [II], and [III] cannot be considered because they have “p”. [VI] cannot be taken into account either because it does not have ‘s’. And, in [IV] and [V], we do not know what the value of “p” is. It can be false but it can also be true.

So, it is evident that it can be noted that the four conclusions do not follow the premises by only paying attention to the canonical models of the latter. This, from MMT point of view, means that checking the defect of the inferences should be a very easy, trivial, and simple task for most of the human beings. Accordingly, it appears that, as claimed, for example, by López-Astorga (2016a; 2016c) with regard to other aspects of Aristotelian logic, this last system can play any kind of role in current cognitive science.

5. Conclusions

True, as said in papers such as that of López-Astorga (2016a), if MMT is a correct framework, certain elements and aspects of Aristotle’s logic have the potential to predict the conclusions that individuals can come to given certain inferential structures. Based on this, it can be thought even that Aristotelian logic (if not in entirety, at least some of its parts) describes the real way people reason. And this is so because, if it is possible to notice just by resorting to canonical models whether or not the conclusion follows in the inferences of Aristotelian logic, it is evident that such inferences correspond to ways people usually reason in everyday life.

This is an idea that is also supported by papers such as those of López-Astorga. However, it is clear that, to be absolutely accepted, as also indicated in works such as the aforementioned ones, empirical evidence is needed. There is no doubt that many experimental results confirm the general assumptions of MMT. However, continuing with the basic approach of those papers, it can be said that empirical studies intended to check arguments such as those presented here could be useful in at least two ways. On the one hand, they would enable a better interpretation and understanding of Aristotelian logic, since they could reveal its actual potential and value. On the other hand, they would also allow checking whether or not certain predictions of MMT (for example, those related to the kind of inferences reviewed in this paper) are correct. Furthermore, the basic design that such studies could have is obvious: keeping the formal

structure of the inferences used by Aristotle, new created versions of them with thematic content would be presented to groups of participants, the task for them being just to indicate whether or not, given the premises, the conclusion must be admitted.

Accordingly, it seems that the line of research opened by the works using MMT to analyze particular aspects of Aristotelian logic deserves to be developed to a greater extent. If, for example, the moods of the four figures correspond to structures that allow deriving correct conclusions easily and the inferences analyzed here match structures with clear incorrect conclusions, it is evident that such structures can be used as research instruments in different contexts. Thus, as argued, the benefits to be found are real both in the study of the logic provided by Aristotle (it could be checked whether or not his schemata are consistent with the actual cognitive human behavior) and in the review of the validity of MMT (it could be checked whether or not people resolve reasoning tasks only requiring canonical models without difficulties). And, of course, all this independently of the purely philosophical, psychological or linguistic interest that studies of this kind can also have.

References

JOHNSON, F.

2004 «Aristotle's modal syllogisms», in D.M. Gabbay and J. Woods (eds.), *Handbook of the History of Logic, Volume 1. Greek, Indian and Arabic Logic*. Elsevier, Amsterdam, 247-307.

JOHNSON-LAIRD, P.N.

2010 «Against logical form», *Psychologica Belgica* 5, 193-221.

2012 «Inference with mental models», in K. J. Holyoak E R. G. Morrison (eds.), *The Oxford Handbook of Thinking and Reasoning*. Oxford University Press, New York, NY, 134-145.

2015 «How to improve thinking», in R. Wegerif, L. Li and J.C. Kaufman (eds.), *The Routledge International Handbook of Research on Teaching Thinking*. Routledge, New York, NY, 80-91.

JOHNSON-LAIRD, P.N., KHEMLANI, S. AND GOODWIN, G.P.

2015 «Logic, probability, and human reasoning», *Trends in Cognitive Sciences* 19, 201-214.

KHEMLANI, S., LOTSTEIN, M., TRAFTON, J.G. AND JOHNSON-LAIRD, P.N.

2015 «Immediate inferences from quantified assertions», *The Quarterly Journal of Experimental Psychology* 68, 2073-2096.

KHEMLANI, S., ORENES, I. AND JOHNSON-LAIRD, P.N.

2012 «Negation: a theory of its meaning, representation, and inference», *Psychological Review* 109, 646-678.

2014 «The negation of conjunctions, conditionals, and disjunctions», *Acta Psychologica* 151, 1-7.

LÓPEZ-ASTORGA, M.

2016a «Does Aristotelian logic describe human reasoning? Valid syllogisms and canonical models», *Khazar Journal of Humanities and Social Sciences* 19, 5-27.

2016b «Against Philo's interpretation of the conditional: the case of Aristotle's thesis», *Agora. Papeles de Filosofía* 35, 51-64.

2016c «The logical relations of opposition and how to teach them», *Opción* 32, 32-49.

LUKASIEWICZ, J.

1957 *Aristotle's Syllogistic: From the Standpoint of Modern Formal Logic*, Clarendon Press, UK, Oxford.

ORENES, I. AND JOHNSON-LAIRD, P.N.

2012 «Logic, models, and paradoxical inferences», *Mind & Language* 27, 357-377.

PEIRCE, C.S.

1931/1958 *Collected papers of Charles Sanders Peirce*. C. Hartshorne, P. Weiss, and A. Burks (eds.). Harvard University Press, Cambridge, MA.

QUELHAS, A.C., JOHNSON-LAIRD, P.N. AND JUHOS, C.

2010 «The modulation of conditional assertions and its effects on reasoning», *Quarterly Journal of Experimental Psychology* 63, 1716-1739.

RAGNI, M., SONNTAG, T. AND JOHNSON-LAIRD, P.N.

2016 «Spatial conditionals and illusory inferences», *Journal of Cognitive Psychology* 28, 348-365.

PREDICTIVE ENGAGEMENT AND MOTOR INTENTIONALITY

Valeria Bizzari^{*} and Inês Hipolito[°]

^{*}*University of Pisa*, [°]*University of Wollongong*

valeria.bizzari@libero.it; hipolito.ines@gmail.com

Abstract: In this paper we aim to show that motor intentionality, as the underlying ground for social cognition, can be explained through the predictive engagement model. Sensorimotor processes seem to play central roles in social interaction, cognition and language. We question the phenomenological role of the body in social cognition and further investigate a causal neural explanation. We will adopt a different perspective by linking the role of the body and intercorporeality with recent findings in philosophy of neuroscience under the predictive brain hypothesis. In fact, the living body seems to entertain a dialogical and enactive relationship with the surrounding context, as well as with neural circuits actively responding to external stimuli. The body is thus configured as a living organism, and not as a mere biological substratum, offering to phenomenology and empirical sciences further confirmations of the possibility-and need-for a cooperation.

Key Words: neurophenomenology, predictive engagement, *Leib*, action, movement.

1. Introduction

The past 20 years have witnessed a change in focus towards the body in the philosophy of mind, cognitive neuroscience and artificial intelligence. In particular, computational explanations of the mind (Searle 1992; Bechtel and Abrahamsen 1991) have been challenged by embodied cognitive science: our main aim is to offer a contribution to this perspective, specifically emphasizing the fruitfulness of a collaboration between phenomenology and neuroscience. In the Second section, we will describe the importance of the living body and of motor intentionality in perception; while in the Third section we will account for the predictive engagement hypothesis, a neuroscientific theory which seems to be compatible with an enactive and embodied view of perception. The beginnings of this approach date back to Maurice Merleau-Ponty (1962). Interestingly, it was in robotics with Rodney Brooks (1991) that the importance of the body to mental life came into focus again. He aimed to construct perceptual input systems and action output systems, linked by a central system that computed representations. Brooks realized, however, that a disembodied inference-based approach could not succeed without its physical environment. He realized that a certain intelligent behavior could emerge without representational models; it can be produced from interactions between the robot and the environment. He suggested that «the world is its own, best model» (Brooks 1991: 139) since the bodily and the envi-

ronmental structures can do the work previously attributed to internal states. By focusing on real-world, real-time, task-oriented cognition, Brooks work inspired and transformed studies on intelligent behavior and cognitive science.

The Embodied Mind by Varela, Thompson and Rosch (1991) followed with the idea that conscious experience depends largely on our bodies and their interactions with the environment in which the cognitive system is embedded. They focused on consciousness and rejected a reductionist view of the mind. Embodied cognitive science aims to understand the full range of perceptual, cognitive, and motor capacities we possess, as capacities that are dependent upon features of the physical body. Therefore, since cognition deeply depends on aspects of the agent's body other than the brain, the involvement of the body in both sensing and acting is crucial for thought and mental affairs. Varela, Thompson, and Rosch introduced the concept of enaction¹ to present and develop a framework that places strong emphasis on the idea that the experienced world is portrayed and determined by mutual interactions between the physiology of the organism, its sensorimotor circuit and the environment. According to this perspective, the action itself cannot be conceived in a merely formal sense, but seems to be the product of the subject's intentional activity, originally connected with its corporeality.

This approach has been very influential both on embodied cognition and on the phenomenological tradition. In fact, there is a growing realization and acceptance in cognitive science that embodied perception is not just a subsidiary module. Nevertheless, this tendency is very well accepted for «low-level» cognition, but if we take into account «high-level» cognition, it seems that computational cognitivism could be the only viable option (Stewart, Gapenne and Paolo 2014). A major aim of this paper is to show that this impression is quite false, and that the paradigm of enaction has its own and highly distinctive approach to higher-level cognition, such as social cognition. We will attempt to bring forward a possible linkage between the phenomenological understanding of social cognition (Section 2) and its potential neural explanation (Section 3). Finally, we will show how the lived body is fundamental in perceptual and cognitive processes through which the subject relates to the world.

¹ «Enaction» is the idea that cognition is grounded in the sensorimotor dynamics of the interactions between a living organism and its environment. A living organism enacts the world it lives in; its effective, embodied action in the world actually constitutes its perception and thereby grounds its cognition (Gallagher and Allen 2016).

2. *The bodily self and social cognition*

2.1 *Self-consciousness*

Human self consciousness depends on the metarepresentation of mental and bodily states as one's own mental and bodily states. However, first-person perspective taking is not sufficient, it is necessary for human self consciousness. In fact, to assign a first-person perspective is to center one's own multimodal experiential space upon one's own body, and thus operating in an egocentric reference frame. Does this have to always be conscious? According to the phenomenological tradition, all conscious states seem to involve self-awareness (though not in Descartes's terms). The sense of self seems to be an integral and ubiquitous part of our experiential life. Experience, in these terms, seems impossible without the sense of self. According to Zahavi (2015; 2008), a mental state is conscious *if and only if* it is «lived through» in such a way that the subject is immediately and non-objectionally acquainted with it as his own. As argued by Gallese and Sinigaglia (2011), mental states or processes are embodied primarily because of their bodily format. Like a map and a series of sentences might represent the same route with a different format which means mental representations might have partly overlapping contents (e.g. a motor goal, an emotion or sensation), they differ from one another in their representational format. This format contains what a mental representation can represent, that is, when planning and executing a motor act, bodily facts constrains what a mental representation can represent (actions, emotions or sensations involved) and corresponding representations involved in observing someone else performing a given action or experiencing a given sensation. Neurobiological evidence and work on special cognition has showed the fundamental role of action and sensorimotor activity in perception and cognitive processing, through neural development and plasticity. The development of neural circuits in the visual system and acquisition of visuomotor skills critically depends on sensorimotor interactions and active exploration of the environment (Kulick *et al.* 2015; Little and Sommer 2014).

Even adult brains show considerable plasticity of cortical maps that are dependent on action context and attention (Singer 2017; Chapman *et al.* 2015). Musicians, who often show functional changes in their sensorimotor system resulting from action-dependent plasticity (Fiori and Guzzetta 2015; Münte, Altenmüller, and Jäncke 2002), show the same level of plasticity. These results seem to indicate that appropriate action and relevant sensorimotor contingencies appear necessary throughout life to stabilize the functional architecture in respective circuits (Engel 2010). Other highly intriguing finding is that motor and premotor systems are also active during «visual actions» (Urgen and Saygin 2015), like, for instance «mental rotation» of objects (Cona *et al.* 2017; Richter *et al.* 2000). To summarize, from a neurobiological perspective, motor system seems to constitute a fundamental part in perception and active engagement with

the environment, what one now needs to be addressed is how this engagement is lived by the subject.

2.2. Motor intentionality

Phenomenological inspiration is today revolutionizing the role played by the body in cognition, especially emphasizing the relevance of the body in intersubjective understanding and, thus, showing the centrality of the body in the higher processes of perception and knowledge. According to the phenomenological perspective, both the cognitive process and consciousness are simply the product of our being embodied. The body is the means by which the subject may live in the world and differ from inanimate creatures. The living body, in fact, is characterized by being intentionally directed outwards (as a starting point for all kinds of knowledge) and by a self-affection that allows it to be aware of itself regardless of any interaction with the world.² One of the main features of the phenomenological approach is *to consider our corporeality as that which is directly responsible of our understanding of the world:*

According to Merleau-Ponty, in everyday, absorbed, skillful coping, *acting is experienced as a steady flow of skillful activity in response to one's sense of the situation.* Part of that experience is a sense that when one's situation deviates from some optimal body-environment relationship, one's motion takes one closer to that optimum and thereby relieves the "tension" of the deviation. One does not need a goal or intention to act. One's body is simply solicited by the situation to get into equilibrium with it. (Dreyfus and Dreyfus 1996: 11, our italics)

In order to explain the entanglement between the corporeal subject and the world, Merleau-Ponty uses the notion of *operative or motor intentionality*,³ that is a passive and lived intentionality through which the bodies are linked to each other and know each other in a non-tetic, pre-reflexive experience. The main

² This dual characterization, along with the dichotomy *Leib/Körper*, has been a source of inspiration to many authors not directly linked to the phenomenological method, which have been focused on fields of study to which classical phenomenology has never been of interest – neuroscience or, more generally, any field which could be defined as cognitive science. Adopting a phenomenological perspective on knowledge has produced, for example, the identification of specific cognitive schemas that belong to the pre-reflective domain: this is how Gallagher, spokesperson of this theoretical change, gave birth to the famous distinction between *sense of ownership* and *sense of agency* – those characteristics of consciousness that allow for an immediate and latent self-awareness, as well as movements and actions which are one's own.

³ The philosopher referenced by such approach is mainly Merleau-Ponty, who argues for the inseparability between bodily abilities and consciousness: in other words, our perception of the world depends on the structural aspects of our existence.

feature of this kind of intentionality is its motricity: the subject moves and acts according to a specific *bodily schema*, a concept that the phenomenologist borrows from the psychologist Schilder.

The *body schema* is the structure through which the body acquires not only self awareness as an “I can” (and not simply as an “I think”), but also the ability to move consciously in a given space. Motor intentionality makes, thus, the subject capable of perceiving and understanding the facticity of self and of the world in which he/she is immersed, and aware of his/her capacities. The being-at-world presupposes a body that moves and perceives not just through one of its senses (such as sight or touch), but with the totality of its being. The reason behind this assumption is that each part of the body is involved in one another, developing an integral and complex experience. What emerges seems to be a complex subjectivity, capable of dealing with the world through a practical knowledge (*praktognosia*) and not merely through mental representations.

The role of the lived body would seem, indeed, to structure the perceptive experience and make it meaningful. This is evident, for example, in the acquisition of a new habit: this does not seem at all to be the result of a purely intellectual operation that takes place by means of representations or inferences, but rather a pre-reflexive, involuntary and corporeal act. The key features of the “habitual” body are the physiological structure (because it depends on the operation of various sensory organs) and the need to be put into practice so that it can preserve its “habitual” nature. The perceiving consciousness and the physiological substrate consciousness cannot be separated because they are dialogically and mutually connected. In other words, the Merleau-Pontian conception of perception involves a circular mechanism necessarily involving the body, the cognitive process and consciousness. The meeting between the body and the world also implies a dynamic and dialectic relationship. Perception, in fact, does not seem to be a mere representation because the «habitual body» is constantly modified by its interaction with the environment: learning to dance or to play an instrument require a change of affordances and of the intentional relationship between the subject and the world. As the very core of perceptual activity, the body is the instrument thanks to which there can be a link between the subject and the world, the ego and the alter ego.

Within this perspective, motor intentionality seems to be a pre-reflective and lived kind of knowledge that allows the subject to recognize the other in an immediate and non-thetic manner. More specifically, concerning the issue of intersubjectivity, Merleau-Ponty claims that the experience of self necessarily presupposes the experience of otherness: essential to the subject is his ontological openness and the tendency of overcoming himself. Furthermore, either in the case of the perception of self, or in the case of the perception of otherness, what is at stake is an *embodied subjectivity* whose main feature is the involving of an alterity. In fact, in the case of bilaterality of kinesthesia, the

subject makes an internal and an external experience, because he/she anticipates both the manner in which the other would experience him/her and the way through which he would have experienced the other.

In this view, our embodied Self-awareness could be described as a pre-feeling of otherness, and the intersubjective experience as an echo of our own corporeal constitution:

My right hand was present at the advent of my left hand's active sense of touch. It is no different fashion that other's body becomes animate before me when I shake another man's hand or just look at him. In learning that my body is a "perceiving thing"... , *I prepared myself for understanding that there are other animalia and possibly other men.* (Merleau-Ponty 1962: 212)

In *Phenomenology of Perception* (1962), Merleau-Ponty explicitly faces this issue (especially in the chapter *The Other and the Human World*) and offers a contribution in the course *Les relations avec autrui chez l'enfant*, where he analyzes the psychoanalytic perspective and some developmental theories according to which we can talk about the perception of the alterity in psychogenetic terms.

According to Merleau-Ponty, we are intersubjective creatures from birth simply because we possess a corporeal schema. It is very interesting to notice that an experimental study conducted by Meltzoff and Moore has shown that newborns (the "youngest" 42 minutes old, the "oldest" 72 hours old) are able to imitate facial expressions, thanks to an inner capability very similar to the Merleau-Pontian corporeal schema which creates a bridge between interiority and exteriority. Without the intervention of simulations or inferential capabilities, the subject is able to perceive the other's corporeal movements as *expressive and intentional* starting from the first year of life, and can immediately understand the other as an *agent*, and not as an object or Cartesian mind.

According to this view, there are no epistemological functions exclusively committed to the understanding of other minds: the subject intuitively understands the rage in the other's gestures or facial expressions. To quote Gallagher: «Such perceptions give the infant, by the end of the first year of life, a non-mentalistic, perceptually based *embodied understanding* of the intentions and dispositions of other persons» (Gallagher 2008: 540). The body seems to be the place of the emergence of (shared) meaning.

3. *Movement and cognition: towards a new neurophenomenological paradigm*

Considering this image of perception makes it possible to argue that the living body is a key element of the whole cognitive process, starting from which you

need to reframe the relationship between cognition and consciousness and, therefore, to think of an alternative embodied approach. In other words, we can affirm that «embodiment does cognitive work» (Morris 2010: 236) because the living body is essentially «a cognizing agency» (*Ibidem*). Within this proposal, *movement* takes on a key role, as it constitutes the main instrument through which we form cognition: through body movements, in fact, the subject *explores* the world, perceives his affordances and determines his habits (*habit body*). In other words, the cognitive process seems to be the result of an essentially embodied perception, which arises from a living body: «At their most fundamental level, subjective experiences are tactile-kinesthetic experiences. They are experiences of one's own body and body movement; they are experiences of animate form. These experiences are the bedrock of thinking» (Sheets-Johnstone 1999: 435). It is possible to trace an emphasis on the role of the body within the cognitive process even in Husserl's thought: every act of perception (*Noesis*) implies a noetic horizon, which makes it possible to have a complete knowledge of the perceived object. In other words, perception is not a univocal experience, since it implies the possibility of assuming different perspectives. For this reason, the ego seems to be a necessarily *moving body*: in fact, differently, it would not be possible to postulate the existence of other horizons of perception and of other points of view. To summarize, we argue that the possession of kinesthetic skills is the characteristic which constitutes the subject as an *animated organism*. Husserl repeats this expression several times, especially in the *Cartesian Meditations*, where it becomes clear how the subject is essentially a psychophysical organism that interacts with the environment.

Phenomenologists and scientists are increasingly focusing their attention on *body movement*: according to Edelman, for example, movement is instrumental to the knowledge of the world, while the American philosopher M. Sheets-Johnstone argues explicitly that «cognition is not separated from perception, perception is not separated from an environment nor from a larger category designated as a behavior: on the contrary, the movement-perceptual system *is* behavior in the sense that it is the actual “real-time”, “real-life” event as it unfolds» (Sheets-Johnstone 1999: 218). In other words, the cognitive process seems to have *affective* features: the movement seems to be the first communication resource, a source of non-linguistic and kinetic concepts (space, time, force, etc.).⁴ Motor activities, as well as emotional experiences, are primary resources for the knowledge of the world, before the arising of more complex cognitive abilities, whose proper functioning seems rather to derive from them.

⁴ M. Sheets-Johnstone provides the example of walking, an essentially qualitative experience that must be phenomenologically described through terms that refer to the subject's kinesthesia.

In particular, the notion of “living body” is useful because implies a psychophysical organism that, by means of its capabilities, and, therefore, movement, not only manages to gain experience of the environment, but also of himself (self-construction of the living body). The link between the subject and the world is therefore a reality shaped by «non-linguistic corporeal concepts» (Sheets-Johstone 1999: 153). In addition, the same neurobiological findings confirm this hypothesis. In fact, it has been shown that during perception sensory-motor interactions actually occur: for example, the development of neural circuits related to the visual system and its capabilities are dependent on the active exploration and interaction between the organism and the environment (Held 1965; Majewska, Newton and Sur 2006). Therefore, body action is not always complementary to cognition, but it is possible to argue that it is itself cognition.

In this perspective, also intersubjective perception⁵ could be outlined as an interactive process, and not purely a cognitive one: as in the encounter with the other, the subject is not a mere observer, but «responding in an embodied way» (Gallagher 2008: 540). In this context, social cognition appears to become a synonym for social interaction, namely a process in which body’s movements, expressions and context play a key role. The idea behind this approach is that our intersubjectivity is essentially *a direct bodily mechanism*, not only during childhood but also into adulthood. A representational account of cognition does not seem, thus, sufficient to explain our mental life (Lakoff and Johnson 1999): on the contrary, we should consider perception as an *active bodily interaction with others and the world*. This perspective seems to be coherent with the *predictive engagement* approach, suggesting a promising direction of study. In fact, bodily interaction has been recently considered within its relation to the free energy principle, in which predictive engagement plays an important role.

3.1 *Free energy principle and predictive engagement*

Predictive modeling uses statistics to predict outcomes in the future or an unknown event in the past. In most cases, the model is chosen on the basis of detection theory to try to predict the probability of an outcome given a set amount of input data. Models can use one or more classifiers in trying to determine the probability of a set of data belonging to another set. Predictive brain is, thus, the

⁵ According to Merleau-Ponty’s perspective, the other plays a crucial role, since intersubjectivity is based on our own identification with a partner’s body, identification which happens through an immediate perceptual linkage with his or her body. Thus embodiment constitutes a necessary condition for the possibility of intersubjective agreement about the shared world.

view that the nervous system maintains internal probabilistic models that are updated by neural processing of sensory information using methods approximating those of Bayesian probability (Hohwy 2013).

There can be a fundamental distinction between three philosophical views on neuroscience of predictive models:⁶ *predictive coding* is associated with internal Bayesian models and prediction error minimization; *predictive processing* is associated with “simple” embodiment; and *predictive engagement* appears associated with enactivist approaches to cognition (Gallagher and Allen 2016; Kirchhoff 2016; Allen and Friston 2016).

On the enactivist model, social cognitive process is seen as closer to ongoing predictive engagement, that is, a dynamical adjustment in which the brain, as part of and along with the larger organism, actively responds in ways that allow for the right kind of ongoing attunement with the environment: an environment that is physical but also social and cultural. Enactivists suggest that the brain is not located at the center, as neural accommodation occurs via constant reciprocal interaction between the brain and body. So, adjustment and attunement can be cashed out in terms of physical dynamical processes that involve brain and body, including autonomic and peripheral nervous systems. We can see how this enactivist interpretation can work by exploring a more basic conception operating in these predictive models, namely, the free energy principle (Allen and Friston 2016). This concept, originally introduced by Karl Friston, has been recently expanded to serve as a bridge between information theory and biological processes.

The free energy principle, also known as active inference, suggests that (biological) systems maintain order by restricting themselves to a limited number of states. In order to do so, systems minimize a free energy functional of their internal states, which entails beliefs about hidden states in their environment.

Furthermore, free energy principle argues that biological systems are foremost defined by the tendency to resist the second law of thermodynamics, on the basis that to do otherwise would entail the unbounded increase of entropy. This models reveals similarities with the autopoietic system by Francisco Varela (1974), particularly focusing on the linkages between embodiment, environment and brain processes through the application of information theoretical principles. According to this model, an organism both generates internal dynamics of probabilistic predictions embodied in neural networks that maximize survival (minimize free energy), and acts on the world in such a way as to cause sensory information to conform to prior predictions (Gallagher and Allan 2016: 9). Fur-

⁶ For a detailed account see the special issue «Predictive Brains», *Synthese* (2016).

thermore, the organism, as an autonomous agent, avoids unexpected states, that is, minimizes surprise.

In short, the free energy principle is an attempt to explain the structure and function of the brain as a minimizer of variational bound on disorder, and it asserts that any adaptive change in the brain will minimize free-energy. This minimization could be over evolutionary time (during natural selection) or milliseconds (during perceptual synthesis). In fact, the principle applies to any biological system that resists a tendency to disorder, from single-cell organisms to social networks (Friston 2009). Free energy is an information theory quantity that bounds the evidence for a model of sensory inputs and the model is encoded by the brain. Free-energy is thus greater than the negative log-evidence or “surprise” in sensory data.

Within the three approaches to predictive model, for *predictive coding* and *predictive processing*, active inference is part of a process that produces sensory experiences that confirm or test expectations, while for *predictive engagement*, active inference is more action than inference; it’s doing a doing, an enactive adjustment, a worldly engagement with anticipatory and corrective aspects already included). It is a loop that also navigates through the body and environment and forma a whole (Gallagher and Allan 2016: 9). However, despite the distinctions, there seems to be some common ground in respect to prior experience and how the system embodies interaction and sociality in the world.

This seems to be coherent with the phenomenological account, according to which, as we have previously described, cognition is dynamically incorporated and located in the environment. Predictive engagement approach can integrate phenomenology in a mutually informed manner, enabling a description of the subject not as an “I think” but as an “I move”.

4. *Conclusions*

In this paper we have considered the hypothesis that social shared world is portrayed and determined by mutual interactions between the physiology of the organism, its phenomenological experience and the environment. The phenomenologically inspired work on cognition is today increasingly influential. There is a growing realization and acceptance in cognitive science that embodied perception is not just a subsidiary module. We aimed at contributing to an alternative model to high-level cognition and hierarchical neural exchange through the *predictive engagement model*. We looked hence into the phenomenological understanding of perception (Section 2) and its potential neural explanation (Section 3.1).

In Section 2 we attempted to explain the phenomenology of the role played by the body in the shared world. We first explained how motor system appears to constitute a fundamental part in perception and active engagement with the

environment, and how this engagement is lived by the conscious self. To further explain this linkage, we focused on a novel viewpoint on intercorporeality: motor intentionality, which seems to yield the subject the ability of perceiving and understanding the facticity of self, and of the world in which he/she is immersed, and aware of his/her capacities. Finally on Section 3, we hope we were able to accomplish the aim of bringing forth predictive engagement as a philosophical position in the neuroscience of predictive brain hypothesis as a liable explanation for social cognition and human sociality. Social cognitive process, in this model, is considered as closer to ongoing predictive engagement, that is, a dynamical adjustment in which the brain, as part of and along with the larger organism, actively responds in ways that allow for the right kind of ongoing attunement with the environment. This brings us to the possibility of exploring a more basic conception operating in these predictive models, namely, the free energy principle. The application of information theoretical principles further permits us to focusing on the linkages among embodiment, environment and brain processes. This, ultimately, seems to leave the door open to look into a new causal explanation for higher cognitive processing, such as social cognition, in its *embodied and phenomenological nature*. The organism both generates internal dynamics of probabilistic predictions embodied in neural networks that maximize survival (minimize free energy), and acts on the world in such a way as to cause sensory information to conform to prior predictions.

Acknowledgments. The authors would like to thank the anonymous reviewers for their mostly insightful suggestions. Inês Hipólito would like also to thank the support of the Lisbon Mind and Cognition Group.

References

- ALLEN, M. AND FRISTON, K.
2016 «From cognitivism to autopoiesis: Towards a computational framework for the embodied mind», *Synthese*, doi:10.1007/s11229-016-1288-5.
- BECHTEL, W. AND ABRAHAMSEN, A.
1991 *Connectionism and the Mind. Parallel Processing, Dynamics, and Evolution in Networks*, Blackwell Basil, Oxford.
- BROOKS, R.A.
1991 «Intelligence without representation», *Artificial intelligence* 47, 139-159.
- CHAPMAN, S. B., ASLAN, S., SPENCE, J. S., HART, J. J., BARTZ, E. K., DIDEHBANI, N. AND LU, H.
2015 «Neural mechanisms of brain plasticity with complex cognitive training in healthy seniors». *Cerebral cortex*, 25(2), 396-405.

CONA, G., MARINO, G. AND SEMENZA, C.

2017 «TMS of supplementary motor area (SMA) facilitates mental rotation performance: Evidence for sequence processing in SMA», *NeuroImage*, forthcoming.

DREYFUS, A. AND DREYFUS, S.

1996 «The challenge of Merleau-Ponty's phenomenology of embodiment for cognitive science», in G. Weiss and F. Harber, *Perspectives on Embodiment: the Intersections of Nature and Culture*, Routledge, New York, 103-120.

ENGEL, A.K.

2010 «Directive minds: how dynamics shapes cognition», in J. Stuart, O. Gapenne and A. Di Paolo (eds.), *Enaction: Towards a New Paradigm for Cognitive Science*, MIT Press, Cambridge, MA, 219-243.

FIORI, S. AND GUZZETTA, A.

2015 «Plasticity following early-life brain injury: Insights from quantitative MRI», in *Seminars in Perinatology* 39, 141-146.

GALLAGHER, S.

2008 «Direct Perception in the Intersubjective Context», *Consciousness and Cognition* 17, 535-543.

GALLAGHER, S. AND ALLEN, M.

2016 «Active inference, enactivism and the hermeneutics of social cognition», *Synthese* 1, 1-22.

GALLESE, V. AND SINIGAGLIA, C.

2011 «What is so special about embodied simulation?» *Trends in Cognitive Sciences*, 15, 512-519.

HELD, R.

1965 «Plasticity in sensory-motor systems», *Scientific American* 213, 84-94.

HOHWY, J.

2013 *The predictive mind*. Oxford University Press, Oxford.

KULICK, J., OTTE, S., & TOUSSAINT, M.

2015 «Active exploration of joint dependency structures», *Proceedings of the IEEE International Conference on Robotics and Automation*, 2598-2604.

LAKOFF, J. AND JOHNSON, M.

1999 *Philosophy in the Flesh. The Embodied Mind and its challenge to western thought*, Basic Books, New York.

LITTLE, D.Y. AND SOMMER, F.T.

2014 «Learning and exploration in action-perception loops. Closing the Loop Around Neural Systems», *Frontiers in Neural Circuits* 7, 7-37.

MAJEWSKA, A. K., NEWTON, J. R., & SUR, M.

(2006) «Remodeling of synaptic structure in sensory cortical areas in vivo», *The Journal of neuroscience*, 26(11), 3021-3029.

MERLEAU-PONTY, M.

1962 *Phénoménologie de la perception* (1945), Gallimard, Paris; trans. C. Smith, *Phenomenology of Perception*, Routledge & Kegan Paul, London.

MORRIS, D.

2010 «Empirical and Phenomenological Studies on Embodied Cognition», in S. Gallagher and D. Schmicking (eds.), *The Handbook of Phenomenology and Cognitive Science*, Springer, Verlag, 235-252.

MÜNTE, T. F., ALTENMÜLLER, E. AND JÄNCKE, L.

2002 «The musician's brain as a model of neuroplasticity», *Nature Reviews Neuroscience* 36, 473-478.

RICHTER, W., SOMORJAI, R., SUMMERS, R., JARMASZ, M., MENON, R. S., GATI, J. S. AND KIM, S.

2000 «Motor area activity during mental rotation studied by time-resolved single-trial fMRI», *Journal of Cognitive Neuroscience* 12, 310-320.

SEARLE, J. R.

1992 *The rediscovery of the mind*, MIT press, Cambridge, MA.

SHEETS-JOHNSTONE, M.

1999 *The primacy of movement*, John Benjamins, Amsterdam.

SINGER, W.

2017 «Formative childhood from a neurobiological perspective», in A.M. Battro, P. Sánchez Sordo and J. von Braun (eds.), *Children and Sustainable Development*, Springer, Verlag, pp. 215-224.

SINGER, T., SEYMOUR, B., O'DOHERTY, J., KAUBE, H., DOLAN, R. J. AND FRITH, C. D.

2004 «Empathy for pain involves the affective but not sensory components of pain», *Science* 303, 1157-1162.

STEWART, J., GAPENNE, O. AND DI PAOLO, A.

2014 *Enaction: toward a new paradigm for cognitive science*, Bradford Book, Cambridge, MA.

URGEN, B. AND SAYGIN, A.

2015 «Representational similarity analysis of fMRI responses in brain areas involved in visual action processing», *Journal of Vision* 15, 503-503.

VARELA, F., MATURANA, H. R. AND URIBE, R.

1974 «Autopoiesis: the organization of living systems, its characterization and a model», *Biosystems* 5, 187-196.

VARELA, F., THOMPSON, E. AND ROSCH, E.

1991 *The embodied mind. Cognitive science and human experience*. MIT Press, Cambridge, MA.

ZAHAVI, D.

2008 *Subjectivity and selfhood: Investigating the first-person perspective*. MIT press, Cambridge, MA.

2015 «You, me, and we: The sharing of emotional experiences», *Journal of Consciousness Studies* 22, 84-101.