Extended Perception

Hsi-wen Daniel Liu (hwliu@pu.edu.tw)
Division of Humanities, Providence University
Shalu, Taichung County 433, TAIWAN

Abstract

The extended mind thesis (in brief, EMT) does not reserve a position for inner representations, which certainly play a substantive role in the making of human mind, however. It is, as a consequence, hard to conceive of perception as extended. The present paper, as a complement to EMT, argues that perceptual processes are extended, and so is perceptual content. The argument proceeds with three reasons. Firstly, perceptual content is action-oriented, supporting perceptual-motor coordination. Secondly, the spatial structure of perceptual content is set externally. Finally, active perception manifests the interactive and selective nature of perception. This complementary work provides a ground for rebutting two recent criticisms and a claimed corollary of EMT.

Introduction

The extended mind thesis, also known as active externalism, maintains that the relation of external resources to an agent is active because the two sides of that relation become a sort of tight coupling (Clark 1997ab, 2002, 2003, 2004, 2005; Clark and Chalmers 1998; Dennett 1996, 1998, 2000). Those external resources are considered as playing a crucial role in “driving the cognitive process in the here-and-now” (Clark and Chalmers 1998: 11). The active processes of managing those external resources, if turning up in the head, would equally be regarded as genuinely cognitive. A question would naturally arise: if the mind is extended like the way it is considered in the extended mind thesis, what is the role that perceptual experiences play? They seem to be purely internal, and hence this can be taken as a challenge to the extended mind thesis. The theme of perceptual experiences, perhaps, can be deemed as a “holy territory” of internalism—especially, the representational theory of the mind—and by contrast seems to be a land very hard to conquer for the extended mind thesis. The processes of perceptual experiences, as generally conceived in the causal theory of content, proceed with the inward direction in general, without the need of coalition with tools or environmental conditions. In addition, perception aims to provide contents and this is the end of perception per se, as it seems. Although perceptual contents can be further used for guiding action, the perception qua perception seems not to be a device for that, or for exploiting external tools or environmental conditions. Furthermore, the main function of perception seems to provide contents that report external conditions, without engaging in problem-solving, hence without involvement in external use. Thus, compared to perception, we seem to find that the themes of language and reasoning are relatively easier to explain for EMT. Language, according to Dennett’s (2000) notion of florid representing, is a mind-tool. While reasoning may be understood as internalized technologies such as arithmetic algorithms deployed on a notepad, it seems much harder to put perceptual experiences in the context of extended mind. As a consequence, the theme of perceptual experiences may be taken as a weak point of EMT. Perception seems to be completely internal.

The above weak point is not incorrigible for EMT. The present paper aim to defend EMT at this point, as a complement to EMT, for three reasons: that perceptual content serves to support perceptual-motor coordination, that the spatial structure of perceptual content is set externally, and that active perception manifests the interactive and selective nature of perception.

Extended Perceptual Content

Perceptual-Motor Coordination

As the first of the three reasons, perceptual content serves to support perceptual-motor coordination. The visual content that a cup is on the table, for example, is not exclusively limited to the faculty of sensation. The visual content is action-oriented as it can be seen as a panel on which various motor routines reside, which manifest the ways in which motor activities can unfold on the top of a visual panel. The handle is a position for cup-user to hold, and the cup-surface requires a different way of holding. Another example, a baseball batter must pay full attention of where and how the pitcher’s ball is arriving. After his repetitive practice, the activity of perceiving-ball-and-hitting-it becomes, as required (after learning), highly integrated. It is a visual-motor activity with subtle varieties, all with smooth integrity, hence in a certain sense becomes a visual-motor merger. The interactions between visual processes and motor processes constitute a two-way channel. Motor processes feed backward, in certain ways, to visual processes (Hurley 1998), which makes vision motor-oriented. This helps the bat-control to respond to the unexpected trajectory of the pitcher’s ball with sufficient flexibility. The motor activities are maintained in a way that they flexibly respond to changeable environmental factors (e.g. unexpected ball-

1 More typically, certain cognitive capacities are taken by internalists as internally privileged, such as self-knowledge, propriopception, introspection, thoughts and second-order beliefs. Such capacities arise internally and are maintained without being determined by external factors.

2 For discussion of similar examples, see Clark’s (1997a) discussion of “affordances”. For example, to humans a chair affords sitting.
trajectory, or even strong wind) that are detected in visual content. Thus, visual content provides various facilities that help a controlling agent to maintain visual-motor coordination and co-operation. Perceptual content, hence, does not end up as perceptual report or private entertainment but provides abundant facilities that indicate various subtle visual-motor combinations. Perceptual content is not independent; rather, it keeps being exploited in support of motor action, through learning in the batter’s repetitive practice and learning from his experiences of competition. The above, in general, manifests a picture of perception-for-use, a picture in which perception is a non-independent component of the perception-motor merger.

Structure of Space

The second reason is that space is an essential element of perceptual experiences, yet it is indeed structured in relation to bodily actions—specifically, our motor ability of reaching with the hands. It is studied in neuroscience that in primates and humans space is represented with the distinction between “far (extrapersonal) space” and “near (peripersonal) space”, the former is the area outside the reach of the hands while the latter being the area that can be reached and grasped by the hands without locomotion (Rizzolatti, Riggio & Shergal, 1994). The representations of space are structured in accordance with the reach of the hands. Note that hands are bodily parts that are deemed as tools by Dennett (2000). Because of space, hence, perceptual contents are distinguished between those which can be integrated with hand activities and those which cannot. As we can infer from the above discussion, the near space can be abundantly exploited by various combinations between subtle representations of the near space and complex hand activities. Representations of far space, by contrast, cannot have such extended combinations.

Moreover, certain parts of the far space can be re-structured by being incorporated into near space, resulting from the expansion of hand control by holding tools like sticks or rakes (Berti and Frassinetti, 2000). The representations of space, in other words, can be re-organized in the way that the area with perceptual-motor combinations expands. The spatial expansion is made through the hand control of holding tools, which brings about abundant complex perceptual-motor combinations. Thus, the spatial structure is re-organized depending on the perceptual-motor combinations that are established outward.

Some skeptics may question whether, in the above consideration of expansion, what is re-structured is perceptual representations but completely motor ones. The use of tools establishes certain motor skills that control areas beyond the finger tips, and thus expands the reach of motor skills. They can even grant that the coverage of perceptual-motor combinations indeed expands, but insist that the perceptual representations of the space per se remain left unchanged. Because motor skills are “filled into” a perceptual space, as those skeptics maintain, what is re-structured is motor space (“the filled staff”), not the perceptual space (“the empty container”).

Such skeptical consideration sounds reasonable. However, it is mistaken because of overlooking the ecological aspect of perceptual representations. The space underlying human perceptual experiences is not simply represented as an empty-container but is full of perceptual-motor combinations that manifest the ways in which motor activities with certain trajectories may arise in certain locations. Such a space is by no means empty before motor activities unfold, but instead full of perceptual guidance of motor activities. While motor activities unfold in space, the motor processes of those unfolding activities are made possible under the guidance of perceptual properties. The agent of motor activities must support/guide her motor activities with perceptual properties, if those motor activities are neither reflexive nor blind. Moreover, the unfolding motor activities also stand as objects of further perception, forming perceptual representations that can be stored in memory for guiding subsequent motor activities. The agent, when holding a stick, can see how motor activities unfold, and feel the changing conditions at the hand location and at the stick location. In other words, the involving motor activities are indeed guided by perceptual representations, as opposed to simply motor ones. The exploitation of space by hand-and-tool has a perceptual aspect. What is re-structured in the expanded near space involves both motor and perceptual representations. Thus, we entrench the notion of re-structuring the expanded near space with perceptual-motor combinations.

Active Perception

The above consideration is limited to lower-level perceptual processes, one may argue, and higher-level perceptual processes remain completely internal. Consider the fact that visual experiences have higher and lower levels, according to Milner and Goodale (1995). One may accordingly contend that vision may be granted as extended but what is extended is limited to lower-level processes. Higher-level vision is different; its processes are completely internal. Against this point, however, the present paper maintains that the higher-level vision remains environment-oriented to a certain degree, although it looks as if it is a device of completely inner manifestation. This point of extended mind is evident in the recognition of active perception, which renders the third reason.

The point of active perception is as follows: perception does not end with re-construction of the external world—providing all details of the world—but instead provides partial information, depending on what the perceptual agent concerns. The notion that vision, in particular, gives rise to a panorama of all details is but an illusion (Nöe 2002). Instead, fine-grained visual content is constructed under the

---

3 Here, space is not taken as something in the internal, but instead an essential form of experience.

4 Such visual-motor combinations may be achieved by learning, and hence are not necessarily Gibsonian affordances.
(selective!) guidance of saccadic attention. Only the limited information needed for the concern, for example guided by a question, will be pursued (Churchland et al. 1994). Visual processes, to consider from an evolutionary point of view, must respond in real-time in order to respond to an encountered environmental challenge, for example for a rat to question whether the moving shadow showing a prey or a predator. Hence, visual processes, in the construction of fine-grained visual experience, provide only the information needed for the agent to respond to her current concern. As a consequence, the visual processes render only a partial view in the visual experience. Visual content provides something (not everything!) external and skips something else.

To consider the thesis of active vision more specifically, the determination of fine-grained visual content depends on the recurrent interactions between saccadic attention (in which goals are implicit) and environmental conditions. Though given the encountered environmental conditions, the visual content would not end up with a fixed destination, as it remains dependent on the agent’s concern and how attentive control of saccades responds to those environmental conditions. Because of the above attention-environment interactions, the fine-grained visual content is not completely dependent on the encountered environmental conditions (counter to the causal theory of content). The generation of fine-grained visual content, though in a very short time, is a result of three interacting components: agent’s goal, attention, and environmental conditions. The above discussions, to summarize, show that the mind in the making of perceptual experiences is indeed extended.

**Active Externalism of Perceptual Content**

**A Novel Account**

Perception, as is generally conceived of in philosophy of mind, is purely an internal matter, and consequently the notion of extended perceptual content maintained above would sound abrupt and bizarre. Perception is firmly considered as internal, although it is true that perception has external links: perception requires environmental stimuli in the first place, and its information can be used to guide actions. It is generally insisted, in philosophy of mind, that perceptual content *per se*, as opposed to those above links, is individuated purely internally. Perceptual experience arises internally and perceptual content is purely an internal matter. Furthermore, even the way in which perception contributes to action is also determined entirely inside the skin-boundary. Accordingly, it seems to be quite an abrupt idea to see perceptual content as intrinsically extended.

The point of the extended perception thesis discussed above, however, is that the determination of perceptual content must take account of interactions between internal processes and external ones. There is no dispute as to whether perceptual experiences *per se*—consisting of phenomenal characters and perceptual content—reside internally or not. It seems to have been assumed in philosophy of mind that perceptual processes—whatever they be, computationalist or connectionist, or even Gibsonian—also take place internally once they have reached perceptual receptors. It is rarely considered, as yet, that perceptual processes, in its way of determining perceptual content, need to frequently go across the skin-and-skull boundary in order to take account of factors across that boundary. It is at this point that the present account of extended perception insists.

The externalism of content in philosophy of mind—as initiated in Putnam’s discussion of Twin Earth arguing for the notion of *wide content* with the slogan “meaning just ain’t in the head”—generally discussed meanings of words (e.g., “water”) and contents of thoughts or propositional attitudes. By contrast, the issue of whether perceptual content is externalist is generally disregarded.

The three main theories of perceptual content—sense-data theory, intentional theory, and naïve realism (Martin 1994)—discuss how the content of perceptual experience consists of, and leave off the processes that bring about perceptual content. The question is whether those processes would have a bearing on the nature of perceptual content *per se*. To put it more sharply, if the processes that bring about perceptual content are indeed extended, what is the sense in which perceptual content can also be considered as extended? Given the justification of extended processes in the second section, the question is now focused on whether we can have substantive senses in which perceptual content is intrinsically extended. This is the issue of this section.

The present account of extended perception, though it can be regarded as a complement of the general EMT, is not a straightforward implication of it. The EMT generally involves the use of tools, but the existence of perceptual content seems to be independent of the use of any tool. In addition, it is for the need of problem-solving that the mind is deemed as extended, but the perceptual content itself does not seem to be involved in problem-solving. Furthermore, the extended mind thesis does not reserve a position for mental content. In the previous section, all these three points have been specifically put in the context of perceptual content and been demonstrated successfully, and consequently the notion of extended perceptual content is justified. Yet, the way in which perceptual content can be considered as intrinsically extended remains not a straightforward derivation from the already established notion of extended mind. It needs discussion.

**Active Content**

Perceptual content is extended because it is active, which is in at least five senses, as manifested in the second section. Firstly, it is active as it is *action-oriented*. Such content provides various facilities for a controlling agent to maintain visual-motor coordination. In addition, perceptual content is active in the sense that it is *expandable* content. Perceptual content can be exploited by the feed-forward motor processes. Even the near space itself is expandable. Furthermore, visual content is active because the saccadic control compiles a series of fine-grained tiny-areas for the
visual content in a real-time scale, as manifested in active perception. Finally, perceptual content is active because of its immanent purpose. The compiling of fine-grained perceptual content is to support agents’ need/concern. Hence, the perceptual content is not simply a matter of description. It has the immanent purpose of supporting adaptation. Given all the above reasons, perceptual content can be seen as immanently active. Yet, this does not automatically explain that perceptual content is intrinsically extended.

Surely, the processes that give rise to perceptual content are extended across skin-and-skull boundary, as discussed in the previous sub-section. In this sense is it sufficient to claim that perceptual content is extended. However, this only shows, strictly speaking, that what is extended are those processes, not yet the content per se. Nevertheless, this extended nature of perceptual content can be affirmed because of such content’s pertinent ways of external connection. This property of being extended is part and parcel to, though not straightforward proof of, its property of being active. The fine-grained parts of perceptual content can be compiled (guided by the attentive control of saccades), consequently can be altered, according to the involving agent’s run-time need/concern. The near space is expandable, and the involvement of perceptual-motor combinations can be exploited, in a way that they constitute the infrastructure of the bare perceptual content. In addition, it’s being purpose-laden makes perceptual content go beyond the epistemic category into a somewhat teleological one. Thus, this paper presents a fundamental view with regard to the nature of perception, and also a fundamental perspective regarding the nature of extended content. Besides, perceptual content, in contrast to its processes, can be put in the category of externalism, because the perceptual content as above considered is partly determined by external/non-perceptual factors.

In Clark’s (2003) extended mind thesis what makes the mind extended, strictly speaking, is not the cognitive processes that go beyond the skin-and-skull boundary but primarily the agent-and-external-resources merger. The smooth and close interactions between agent and external-resources result in that merger, which is an emerging unity (though heterogeneously including both perceptual content and external factors) dedicated to a certain problem-solving activity. The perceptual content plays an important role in the making of such an across-skin-and-skull merger/unity. To wit, it warrants both the aforementioned property of being active and that of being extended, including the support in favor of an agent’s need in her environment. Without the perceptual content such a warrant would be pointless to arise. The perceptual content, with all its infrastructures such as the visual-motor combinations previously discussed, warrants a unity with the above-mentioned property of being active and extended. The warrant of that unity makes perceptual content extended in a genuine sense. Not only the processes that make perceptual content are extended (in the sense of across-skin-and-skull processes) but also the perceptual content per se is extended (in the sense construed above).

The perceptual content as an across-skin-and-skull merger, as above considered, makes available a transparent perceptual space that bear realities. Realities seem to be presented in space without the mediation of sense-data or a perceptual space. Everything presented in the perceptual content is seen as a real thing as it can touch our bodies and can be captured by hands. This is made possible by that across-skin-and-skull merger, as the perceptual space has the agent’s body as its origin of coordinate and also serves to guide motor activities. A perceptual space, in theory, can be non-transparent like this, however. The space that we perceive in a movie or a TV program is not as transparent as the “real” space we perceive. This can be explained in our account in terms of lacking the aforementioned merger. This is because the movie does not employ the audience as the origin of its perceptual space, and, above all, the audience cannot touch or catch anything in the movie.

**Criticisms**

**Inner Representations**

Sterelny (2004) criticizes EMT by proposing that the use of epistemic tools disproves EMT’s active account of tool use. In EMT, as Sterelny (2004) understands, the relation of external resources to an agent is active, because the two sides of that relation become a sort of coupling. Those external resources are considered as playing a crucial role in “driving the cognitive process in the here-and-now” (Clark and Chalmers 1998: 11). The active role of external factors marks the EMT as active externalism, as opposed to content externalism in philosophy of mind. The content in the content externalism need not be active. When a believer and her duplicate twin have beliefs about water, what makes their beliefs different are the relevant external features that passively locate at the distal end of the causal chain. With the above understanding, Sterelny criticizes Clark’s active externalism. That role (and consequently the agent-tool coupling) would make the tool use limited to a specific end in specific circumstance, but this is not evident in the way humans employ epistemic tools, as Sterelny contends.

Sterelny’s general line of argument is focused on two themes concerning epistemic artefacts, such as recipes, handbooks and field guides. The first theme is that the information involving in the use of external resources is portable, that is, not tightly coupled to a specific task, but can be used as multi-purpose representation. The second theme is that the use of epistemic tools, Sterelny contends, must involve rich resources of inner representations that accurately track world features and accordingly make possible the open-ended use of those tools, for various projects in even unpredictable circumstances.

The examples employed in Sterelny’s argument are appropriately constructed, and the exposition of those two themes is tenable. Despite these, the argument does not consequently amount to a denial of Clark’s EMT. The EMT

---

5 For details, see the “Introduction” in Schantz (2004).
remains standing in a potentially good position to explain everything considered in the above two themes concerning epistemic artefacts. Although Clark rarely discussed the coupling of epistemic tools in his construction of EMT, this does not preempt the EMT from application to such artefacts. Because they are epistemic artefacts, handbooks and field guides stand at a higher level of cognitive use. The here-and-now circumstances of their use involve various epistemic factors, including the identification of objects and the tracking of features relevant to an agent’s concern, also including ecological information, the agent’s interpretation and the relevant knowledge of those tracked features. Despite its involvement of higher level cognition, an epistemic artefact may well constitute a tight couple with their users, once they have coped with its use in a manner that the youth uses cell phone and its accompanying facilities. Computer can be regarded as an epistemic artifact as it can run lots of highly epistemic programs. Only for those who can smoothly manage such programs that an agent-and-computer coupling can be constituted. The involvement of various relevant knowledge and open-ended use does not preempt the agent and the computer from constituting a tight couple, and consequently the agent can have an active use of her computer. The computer belongs to her not in the sense of ownership, but in the sense of that coupling. That tools can be used flexibly for different users in various circumstances does not mean that there cannot be tight (user-and-tool) coupling and active use (of such tools).

Despite the above-mentioned flaws, Sterelny’s (2004) argument is certainly smartly designed, as it presents abundant ways of tool use for problem-solving that requires inner representations. The role of inner representations is worth attention, as they indeed take part in problem-solving and hold an externalist role in their interactions with the environmental conditions. It is, however, a role that EMT rarely touches. Sterelny’s (2004) denial of the active externalism (EMT) is unsuccessful, as above-mentioned; this seems to manifest a hint that inner representations can play a role in active externalism, but this needs justification.

The present paper can serve as an argument in support of that envisaged justification. As discussed in the second section, the combinations between visual content and motor routines are inner representations that serve to control agents’ motor activities. The role of those inner representations in the control is to promote agents’ motor flexibilities in response to even unpredictable environmental challenges. Note that the perceptual-motor coordination clearly manifests a coupling unity of perceptual and motor systems. As limbs can be seen as external resources, according to Dennett (2000), here we find evidence in support of active externalism. In addition, the attentive control of saccades has immanent inner representations, consisting of saccadic movement responding to certain specific visual features that are especially concerned by an agent under a certain query or interest. The visual-motor combinations in the near space are also inner representations that play a role in the unfolding of the extended mind, as discussed in the second section. Thus, the present paper fills a gap between inner representations and active externalism, and consequently can serve to justify that inner representations can play a role in active externalism.

**Non-Derived Content**

Adams & Aizawa (2001) claim that external processes cannot provide non-derived mental representations: “[in]tracranial cognitive processes involve non-derived mental representations, whereas extracranial processes such as rotating blocks on a computer screen, making marks on paper, tying strings around fingers, do not” (Adams & Aizawa 2001: 57). The present paper, in the second section, shows that perceptual content indeed arises from extended processes. No interpretation, as we should note, is involved in the course of content generation. The inner representations involving in the perceptual content, as a consequence, are non-derived. Hence, the extended perceptual content remains completely non-derived.

**The World Leaking into the Mind?**

Dartnall (2005) raises a version of internalism claiming that the world leaks into the mind. The main reason is, as Dartnall puts it, that “we sometimes perform actions in our heads that we usually perform in the world” (p. 142). This internalism has an epistemological implication: “if a process gives us an empirical discovery when it is performed in the world, it will also give us an empirical discovery when it is performed in the head” (Abstract, italics added). Dartnall’s evidence is mental operations such as mental rotation applied to mental imagery. We can manually work a jigsaw puzzle and find certain geometric relations between shapes; correspondingly, we can find similar results by performing similar operations mentally on the basis of mental imagery.

Such an internalist account does not seem to be tenable, however. It is hard to say that the world leaks into the mind, and this would not be a corollary of the active externalism. Dartnall (2005: 139) argues that the use of external objects as memory stores (as our cognitive model, in a Brooksian view) has the deficiency of non-portability, and that having inner analogs of it “would overcome this problem and free cognition from the here and now, the context and the moment”. That argument, on the one hand, underestimates the cognitive role of the world model, in particular the active role of external resources playing in human-tool coupling. The information of external objects’ activities, in addition, can still be maintained whenever it is needed and consequently this information is largely no less portable than the information provided by rotation on visual imagery.

On the other hand, Dartnall’s (2005) claim exaggerates the effectiveness and efficiency of operations on visual imagery. The operations on the inner analogs of objects and states of affairs, as is generally admitted, would be less effective and efficient than those working on external objects.
This is because the limited storage and the nebulosity of mental imagery compared to perception. Despite their being seemingly closer (to the agent’s management of problem-solving) than the external ones, those inner ones are indeed harder to manage. The concrete facilities and the between-parts relational sharpness provided by the active human-tool coupling would be likely to lose its vividness and clarity in mental imagery. Managing arithmetic operations on large numbers would be much harder then that practiced on paper. When we attempt to mentally operate a complex jigsaw, we are liable to un-attentively make mistakes or even completely get lost. Such deficiencies, however, can be significantly reduced in manual operations on real physical objects. Dartnall happily applauds that employing inner analogs of external operations will “free cognition from the here and now, the context and the moment” (italics added). Yet, this indeed is to erode the abundant subtle information provided in the here-and-now operations and their complex interactions. The mind can leak into the world, but the reverse is hardly promising.

This subsection raises a caveat. The active coupling between agents and tools needs to be managed with our body/limbs/hands in the environment; the substituting operations on visual imagery would not be equally active. Although vision is extended as argued above, mental operations on visual imagery has gone off the ground of human-tool coupling, a ground on which the extended mind must rely.

Concluding Remarks
This paper, as a complement to EMT, argues that perceptual content is extended, extended not in the trivial sense that perceptual content links externally to sensory input and behavioral control. Rather, it is extended in a substantive sense: perceptual processes are extended, and so is perceptual content per se. The visual content is itself extended for several reasons. It is action-oriented, as it can be seen as a panel on which various motor routines are tightly bound. The visual space is not an empty container, but has the immanent infrastructure of abundant visual-motor combinations. Furthermore, the visual space can be re-organized after using tools that seem to extend hands. Still further, the determination of visual content depends across skin and skull on the recurrent interactions between attentive control of saccades, agents’ goals, and the environmental conditions. For such reasons, the present paper shows that perceptual content is intrinsically extended and that certain inner representations (which are bearers of perceptual content) can play a role in the extended mind.

In addition, the above reasons can serve as a counterargument against two criticisms against the EMT. Adams & Aizawa (2001) claims that external processes cannot give rise to non-derived content. The present paper shows that perceptual processes indeed give rise to non-derived content although they are external. Sterelny (2004) criticizes that the human-tool coupling would jeopardize the open-ended use of inner representations. The present paper disagrees, showing that the employment of inner representations in the making of visual content indeed serves to enhance its flexibility. Besides, Dartnall (2005) maintains that world leaks into mind. The present paper criticizes it, as visual experience transformed into imagery would lose its active thrust.

Acknowledgments
This research is supported by National Science Council, Taiwan, under grant NSC 94-2411-H-126-003.

References