Investigating the role of movement in the constitution of spatial perception using the Enactive Torch

Lorenzo Grespan* Tom Froese* Ezequiel A. Di Paolo* Anil K. Seth*
Adam Spiers† William Bigge*
(*) Centre for Computational Neuroscience and Robotics, University of Sussex, UK
(†) Bristol Robotics Laboratory, University of Bristol, UK
E-mail: l.grespan@sussex.ac.uk

Abstract

This paper reports an exploratory study designed to clarify whether the Enactive Torch, a custom-built minimalist distance-to-tactile perceptual supplementation device, can be used to investigate the role of embodied action in the perception of external spatiality. By constraining the kind of exploratory movements available to the participants, we create an experimental setup in which it is possible to study the relationship between bodily degrees of freedom and spatial perception. We present a preliminary investigation of the strategies used by minimally trained participants to locate various objects placed in front of them by engaging in active exploration under constrained conditions.

1. Introduction

Since the early 1990s there has been a growing consensus within the cognitive science community that the body shapes the mind [7, 4]. At present, the challenge is to build on this general consensus by further explicating the specific contribution of embodiment to our mental capacities. One particular focus of interest in this respect is the constitutive role of embodied action for perception [10, 11, 15]. The study of this kind of 'enactive' perception is greatly facilitated by the use of novel technological interfaces [9, 2], especially sensory substitution devices (also known as "perceptual supplementation" devices for reasons given in [8]).

Already in 1969 Bach-y-Rita and his colleagues employed a vision-to-tactile system called TVSS, as "a practical aid for the blind and as a means of studying the processing of afferent information in the central nervous system" [17]. They demonstrated that active exploration with a TVSS, which essentially consists of a camera hooked up to an array of tactile stimulators located somewhere on the body, allowed trained blind subjects to perceive the world as if seen through a camera. Moreover, some subjects spontaneously reported the experience of an externalisation of the stimulation on their body into the world that is in many respects similar to vision.

This seminal study opened a vivid debate about the phenomenology of perception enabled by the use of perceptual supplementation devices, which still continues in the cognitive sciences today [6, 13, 10, 3, 16]. However, despite four decades of research into perceptual supplementation devices, as well as a growing fascination with the phenomenological aspects of their usage, so far no consensus has been reached on how to best understand this type of technology. Indeed, there are ongoing disagreements about some of the most fundamental issues, especially in terms of whether the afforded perception is (i) essentially an extension of the substituting perceptual modality, (ii) the constitution of percepts in the substituted modality, or even (iii) the constitution of a new way of perceiving that is dependent on the specific kind of sensorimotor profile provided by the technological interface [1]. This situation is made even worse due to the fact that the proponents of competing theories often cannot even agree on the experiential phenomenon, i.e. what it is like to use a perceptual supplementation device, that is to be explained.

Some preliminary steps toward the development of a pragmatic phenomenological research program that could address these difficulties were reported by Froese and Spiers [6]. They introduced the *Enactive Torch* (ET), a minimalist perceptual supplementation device, precisely for this purpose. Here we complement those efforts by testing whether this device is also a suitable tool for psychological experiments, especially for investigating the role of embodied action in the constitution of spatial perception. In this paper, thus, we shall revisit some some work originally done by Lenay and Steiner [9], by using the ET. The main objective, apart from testing the original results using a different setup, is to put the ET to a more rigorous experimental test in order to identify the advantages and potential problems with this new scientific tool.

2. The Enactive Torch

In response to the lack of agreement about fundamental issues pertaining to perceptual supplementation technology, Froese and Spiers [6] developed the Enactive Torch (ET), a minimalist device that has been designed to be cheap, non-intrusive as well as easy to use. Accordingly, the ET has the potential of becoming a widely distributed research tool within the scientific community, and thereby help to move the seemingly open-ended debate about the nature of perceptual supplementation forward. In particular, its aim is to inform the specification of the phenomenology of using perceptual supplementation devices by more easily giving researchers first-person access to the experiences in question, an essential source of insight that has so far been sorely lacking in this debate, as well as in the cognitive sciences more generally [14]. A second-generation prototype of the ET is shown in Figure 1 below.

Figure 1: A second-generation prototype of the Enactive Torch (courtesy of A. Spiers).

The main body of the ET contains the power source (batteries) and the circuitry; the separate handle is equipped with an ultrasonic sensor mounted on its end, a small servo-motor with a rotating disc and a vibrotactile actuator. The vibro-tactile motor can generate a set of vibration patterns of variable intensity that can be felt by gripping the handle. In its normal mode of operation the strength of vibration/angular displacement of the disc is proportional to the distance of the closest object in the ultrasonic sensor's range.

In this work, the servo-motor is inactive; we only made use of the vibro-tactile output. The ET is employed in 'binary mode', i.e. the strength of the response can only assume all-or-nothing values according to whether or not an object is present in the ET's field sensor range. The maximum range in this mode of operation is limited to approximately 60 cm; objects are detected if localised within a cone of aperture ca. $30°$.

3. The experiment

This study is inspired by the work of Lenay and Steiner [9] who used a minimalist interface to investigate aspects of perceptual awareness. Their interface was composed of a single photo-electric cell that triggers a binary tactile stimulator whenever the incident luminosity within a cone of about $20°$ is greater than a specific threshold value. Even with such a simple device the localisation of luminous targets is still possible through active exploration. Moreover, it was found that the perception of depth requires a greater capacity for action than the detection of a target object's orientation in relation to the body of the participant. The authors thus argue that the space of lived experience is co-extensive with the space of action and perception, and that the perception of objects does not occur separately 'behind' the perceiver's point of view, but rather in the very same space in which the perceiver moves.

We implemented a set of experiments specifically designed to replicate the work by Lenay and Steiner with a novel perceptual supplementation device: the ET was combined with a simple controlled environment which allows participants to explore the experimental setup with 1 and 2 degrees of freedom (DoFs). The purpose of this study is to measure to what extent blindfolded participants can perceive the position of a target in a novel environment. It is organised into two tasks:

Task 1: Participants are asked to detect the horizontal displacement of a target object by moving the ET horizontally along a fixed 1D axis (one DoF).

Task 2: Participants are asked to detect the distance to a target object by a combination of horizontal movement and rotation of the ET about its centre (two DoFs).

The additional degree of freedom in the second task provides a basis for the participants to perceive distance in addition to an object's horizontal displacement.

3.1. Methods and materials

The experimental setup consists of a sliding platform placed on top of a 160 cm long rail (see Figure 2). Objects are placed in the test-space in front of the rail. The ET is mounted onto the platform, thus being constrained to horizontal movement. The participants can move the platform to both ends of the rail by extending their arm and, if needed, sliding along with their chair. The possibility of using the rotational DoF by turning the platform is only enabled for the second task.

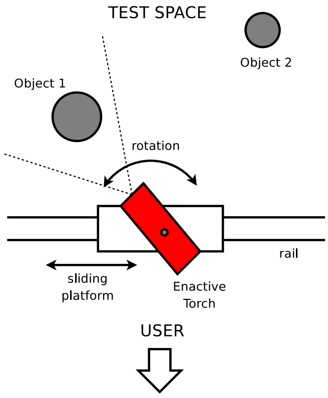

Figure 2: Experimental setup. The user (indicated by the large arrow) can perceive the presence of an object within the ET's range by sliding and rotating the platform on top of which the device is mounted.

The objects are classified according to their size as small (3 cm), medium-sized (9 cm) and large (32 cm), and the distances with respect to the rail are classified as near (8 cm), medium (16 cm) and far (42 cm). We make use of one small and two medium-sized cylindrical objects, as well as one large flat object (a 'wall').

The conic shape of the sensor's receptor field gives rise to an 'inverse shadow' effect: the farther the object, the larger it appears (see Figure 3). This effect could potentially be used by participants as a criterion to distinguish between near and far objects, solely on the basis of their apparent horizontal length.

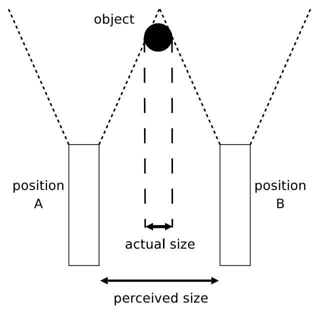

Figure 3: The 'inverse shadow' effect. Assuming the device is moving from left to right, the user will start perceiving the object at position A and stop perceiving it at position B. The object's apparent size is bigger than its actual size.

To prevent participants from perceiving an object's distance as a simple function of its apparent size, we provided them with an experimental situation that cancels out the regularity of the 'inverse shadow' effect through ambiguity. Accordingly, the size of the 'wall' has been chosen such that, when placed at a near distance, it appears as large as a small object placed on the far end. Moreover, the 'wall' is flat to discourage participants from paying attention to its depth, encouraging them instead to distinguish the object only by means of its length (during the second task).

Since the ET was set to operate in binary mode and all the objects were placed within the range of its receptor field, the tactile response triggered by the presence of an object does not indicate that object's distance. Hence, we expect that the minimum number of DoF for which perception of distance can emerge is two.

3.2. Participants

Sixteen participants volunteered to take part in this study, mostly researchers in the fields of Informatics and Psychology. Participant mean age is 31.66 with standard deviation 12.77; there are two outliers (age 65 and 45). In the pool of participants 43.75% (7 participants) are women, 87.5% (14 participants) are right-handed and 31.25% (5 participants) are non-native speakers of English. Two of the participants were already familiar with the ET device; all participants received preliminary training until a basic level of competence was achieved. The experimental tasks were novel to all of the participants.

3.3. Experimental protocol

The number of trials per training task (marked by * in the list below) ranges from three to five, depending on a participant's ability. The participants are blindfolded while carrying out these trials, but they are allowed to visually verify their answers after each trial (this was not allowed during Tasks 1c and 2b to avoid implicit training). At least one trial of Task 2a involved the 'wall' as well as a small object placed at different distances. In this manner we provided participants with the experience that they could not rely on the 'inverse shadow' effect in order to make judgements about distance. The training process allowed them to acquire a sufficiently good mastery of the experimental setup on average within three minutes of first using the ET.

1. Task 1: sliding only (1 DoF)
   (a) * count number of objects (2-4)
   (b) * determine the wider of 2 objects
   (c) determine the centre of 1 object
2. Task 2: sliding and rotation (2 DoFs)
   (a) * determine the further of 2 objects
   (b) evaluate distance to 1 object

For Task 1c participants were asked to explore the target space by moving the ET horizontally, and to place the ET pointing in the (perceived) centre of the object. At this point the experimenter takes a measurement of the correctness, replaces the object, and asks the participant to perform another trial. During Task 2b the participant was asked to declare the object's distance (either 'near' or 'far'). The experimenter records the answer, replaces the object, and asks the participant to perform another trial. The participants remained blindfolded during both tasks, and no form of verbal feedback was given by the experimenter. More details about the objects used in the trials is given in Table 1.

Task	# objs.	obj. size	obj. distance	# trials
1a	2-4	any	any	3-5
1b	2	different	same	3-5
1c	1	any	any	3
2a	2	different	different	3-5
2b	1	any	any	4

Table 1: Experimental parameters. An object size of 'any' indicates that they could have a different or same size; an object distance of 'any' means they were placed at a random distance of 'near', 'medium' or 'far'.

While carrying out the experiments, we took measures to minimize contextual clues - such as the noise produced by placing objects in the test-space - that could potentially be exploited by participants to answer successfully by means other than those intended.

4. Results

Task 1c was achieved successfully by all participants in every trial. Reported centres generally differed by no more than 1 cm from actual centres. Since the ultrasonic sensor is slightly inconsistent across object shapes and textures, an average error of 1 cm is reasonable and we did not feel it necessary to look for greater accuracy.

In terms of Task 2b, the experimental results show that two DoFs are sufficient for the participants to detect the distance of target objects (see Table 2). Participants correctly classified the object distance as 'near' or 'far' in 81.25% of the cases (standard deviation is 0.2627). Those participants who reported using a specific strategy to solve the task correctly classified distance in 88.1% of the cases (standard deviation is 0.1597).

Most participants reported attempts to generate a strategy to carry out the task at hand, in particular with respect to the more elaborate Tasks 2a and 2b. We distinguished these reported strategies into three categories: (i) *cognitive*, (ii) *intuitive*, or (iii) *unknown*. In category (i) we placed all approaches that are based on some explicit geometric/analytical thinking. In these cases, once a strategy has been developed, the participant generally tries to carry it out as if performing the steps specified by an algorithmic procedure. In category (ii) we placed those approaches that rely on some kind of intuitive feelings or pre-reflective bodily skills. Here, the participants judged the success of their embodied actions in terms of a felt sensation. The last category (iii) includes those participants that reported not being aware of any way of solving the task, and who thus resorted to guessing. Table 3 shows the results for each of these categories in terms of Task 2b.

participants	accuracy	std. dev
all	81.25%	0.26%
with strategy	88.10%	0.16%

Table 2: Results of Task 2b. The figures in the first row include two participants who reported not having developed any way to solve the task.

Category	# participants	accuracy	std. dev
cognitive	11 (68.75%)	90.91%	0.15
intuitive	3 (18.75%)	77.78%	0.16
unknown	2 (12.5%)	33.34%	0.33

Table 3: Results of Task 2b for different categories of behavioural strategies. The cognitive strategies are significantly better than the intuitive strategies.

Most participants (68.75%) reported to have used a cognitive strategy, and this strategy turned out be significantly better when compared to the intuitive strategies. Note that the two participants in the 'unknown' category performed worse than chance level (50%), though this is likely due to the small number of trials. We will now describe the behaviour involved in the cognitive and intuitive strategies in more detail.

Cognitive strategies. During Task 1c, when the ET was limited to only sliding along the rail, there was only one type of strategy that was reported. First, the participant would explore the space, until the object was detected. Second, the width of the object is scanned by slowly moving the platform at a constant speed. This provides the participant with a rough estimate of the length of stimulation experienced while the device 'traverses' the whole object. Finally, the participant backs up for half the length of stimulation and then stops. This should leave the ET pointing near the centre of the object. Although only four participants explicitly reported having adopted this method, it has been observed in others as well.

With respect to body movement/posture, it is interesting to note that all participants tried to minimize the number of moving parts: three subjects were observed to only move the arm and keep the rest of the body still, whereas three others kept the arm fixed in position and moved the body instead (by sliding their chair horizontally). One participant reported using the elbow as a marker for the location of an object's earliest detection since the first training session (Task 1a); the marker was later exploited to yield an estimate of the object's size, as per Task 1b.

The extra DoF in Tasks 2a and 2b enabled a broader range of strategies in comparison to the approaches developed for the 1 DoF tasks. The most frequent strategy (observed in seven participants) consisted in (1) pointing the ET at the centre of the target object, and (2) rotating it in both directions until the object was out of the sensor's range. While behaviour (1) essentially consisted in the strategy reported for Task 1c, behaviour (2) made it possible to detect the distance of an object because nearby objects would generate longer stimulation during rotation than far objects.

Another approach relied on the 'inverse shadow' effect to give an estimate of distance as a function of perceived size. Even though all subjects were aware that this regularity holds only for same-shape objects, four participants reported using it as an initial estimate, switching then to the rotational strategy as a method for validation. In particular, by rotating the ET participants were able to detect if the target was the 'wall' - and if not so, then the inverse shadow effect can be thought to hold to some extent. When the target object was indeed the wall, the task took significantly longer to complete, albeit consistently with the correct answer (i.e. it is a 'near' object).

Another interesting example comes from a participant whose approach in Task 2a consisted in (a) positioning the ET about halfway between the two objects, then (b) rotating the handle until the first target object was out of range, and finally (c) repeating (b) for the second object. The extent of rotation needed for stimulation to cease was used as an inverse correlate of object distance, and the objects could be related to each other as 'near' or 'far'. In Task 2b this participant used an approach which consisted in sliding the ET horizontally in one direction until it reached the end of the object; at this point the ET would be slid and rotated in the opposite direction, tracing out tangents to the object. The rate of decrease of the angle was used as a source of information to determine the object's distance.

Finally, one participant devised a strategy for Task 2b based on positioning the ET near one of the ends of the sliding rail. The ET is then rotated, up to a 90° angle, until the device detects the object. If empty space is detected between the object and the rail, the object is reported as being 'far'.

Intuitive strategies. Three participants reported using some kind of "feeling" in order to achieve the tasks, and were unable to provide a step-by-step description of their behavioural strategy. They reported finding it difficult to move the device at a constant speed during Tasks 1b and 1c, as well as having to rely on "sensations" to estimate the width of objects. Interestingly, one of them reported visualising an "imaginary space" for this task.

Regarding Tasks 2a and 2b, the three participants stated that they were unable to find an explicit strategy for distance estimation, and thus decided to rely on the intuitive sense of distance that was generated through their exploratory actions rather than on some form of analytical thinking.

Note that significantly less participants adopted an 'intuitive' strategy. Considering that under non-experimental circumstances, such as when using the ET to find your way across a room, most people report having an intuitive sense of their spatial environment within minutes of their first exploratory activities, this is slightly odd. We speculate that this discrepancy could be the result of a more general problem, namely the attempt to study enactive perception under controlled and minimalist conditions, even though such unnaturally constrained situations are more likely to elicit detached problem-solving attitudes in the participants. If this is indeed the case, then this field would be confronted with even more significant methodological problems than previously assumed.

5. Discussion

It is worth noting that some of the participants spontaneously made observations about their experience that could be interpreted as indicating the beginnings of the constitution of spatial perception. For example, during Task 2b one participant reported that using the ET for active exploration with two DoFs was like "projecting my consciousness forward", while another said that it "felt like being able to see around the object". It appears that, at least for these two participants, their activity of using the ET might have indeed resulted in them being intentionally directed into the world through an actively constituted perception of space.

Interestingly, and contrary to our expectations, these two participants also reported using a behavioural strategy that is based on cognitive inference rather than on skilful tool manipulation. However, it is possible that they conceptualised their pre-reflective behaviour in this manner after being prompted to explain their actions by the experimenter (cf. [5]). In future work it will be essential to record the actual movements of the participants in order to be able to classify their strategies according to objective movement, and not just verbal reports. Such a study would also benefit from better means of collecting phenomenological reports, perhaps by means of second-person interview methods (e.g. [12]). Accordingly, the analysis of strategies presented here should only be viewed as preliminary.

Future work could use the ET to investigate Dreyfus and Dreyfus's [5] description of a progression from novice to expert in a task. While we typically follow rules when we are getting used to a novel task, we become less reliant on such rules as we become more skilful. If behavioural strategies could be objectively classified into cognitive and intuitive, then this suggests a future study where the developmental progression of one class of strategies into the other could be investigated.

Acknowledgements

The authors are grateful to the participants who volunteered to take part in the experiment. They would also like to extend their thanks to all those who contributed to building the experimental scenario and refining its concept, as well as to Nele Froese for providing extensive feedback on a draft of this paper.

References

[1] M. Auvray and E. Myin. Perception with compensatory devices: From sensory substitution to sensorimotor extension. *Cognitive Science*, in press.

[2] M. Auvray, D. Philipona, J. K. O'Regan, and C. Spence. The perception of space and form recognition in a simulated environment: The case of minimalist sensory-substitution devices. *Perception*, 36(12):1736–1751, 2007.

[3] N. Block. Tactile sensation via spatial perception. *Trends in Cognitive Sciences*, 7(7):285–286, 2003.

[4] A. Clark. *Being there: Putting brain, body, and world together again*. MIT Press, Cambridge, MA, 1997.

[5] H. L. Dreyfus and S. E. Dreyfus. Peripheral vision: Expertise in real world contexts. *Organization Studies*, 26(5):779–792, 2005.

[6] T. Froese and A. Spiers. Toward a phenomenological pragmatics of enactive perception. In *Proc. of the 4th Int. Conf. on Enactive Interfaces*, pages 105–108, Grenoble, France, 2007. Association ACROE.

[7] S. Gallagher. *How the body shapes the mind*. Oxford University Press, New York, NY, 2005.

[8] C. Lenay, O. Gapenne, S. Hanneton, C. Marque, and C. Genouëlle. Sensory substitution: Limits and perspectives. In Y. Hatwell, A. Streri, and E. Gentaz, editors, *Touching for Knowing: Cognitive psychology of haptic manual perception*, pages 275–292. John Benjamins, Amsterdam, The Netherlands, 2003.

[9] C. Lenay and P. Steiner. Externalism and enaction. In *Proc. of the 4th Int. Conf. on Enactive Interfaces*, pages 145–148, Grenoble, France, 2007. Association ACROE.

[10] A. Noë. *Action in Perception*. MIT Press, Cambridge, MA, 2004.

[11] J. K. O'Regan and A. Noë. A sensorimotor account of vision and visual consciousness. *Behavioral and Brain Sciences*, 24(5):939–1031, 2001.

[12] C. Petitmengin. Describing one's subjective experience in the second person: An interview method for the science of consciousness. *Phenomenology and the Cognitive Sciences*, 5(3-4):229–269, 2006.

[13] J. Prinz. Putting the brakes on enactive perception. *Psyche*, 12(1):1–19, 2006.

[14] F. J. Varela and J. Shear. First-person methodologies: What, why, how? *Journal of Consciousness Studies*, 6(2-3):1–14, 1999.

[15] F. J. Varela, E. Thompson, and E. Rosch. *The embodied mind: Cognitive science and human experience*. MIT Press, Cambridge, MA, 1991.

[16] P. B. y Rita. Sensory substitution and qualia. In A. Noë and E. Thompson, editors, *Vision and Mind: Selected Readings in the Philosophy of Perception*, pages 497–514. MIT Press, Cambridge, MA, 2002.

[17] P. B. y Rita, C. C. Collins, F. A. Saunders, B. White, and L. Scadden. Vision substitution by tactile image projection. *Nature*, 221(5184):963–964, 1969.