# Bee-ing In the World: Phenomenology, Cognitive Science, and Interactivity in a Novel Insect-Tracking Task


**Guilherme Sanches de Oliveira (gui.cogsci@gmail.com)**
Department of Philosophy & Center for Cognition, Action and Perception, University of Cincinnati
2700 Campus Way, McMicken Hall, Cincinnati OH 45221

**Christopher Riehm (riehmcd@mail.uc.edu)**
Department of Psychology & Center for Cognition, Action and Perception, University of Cincinnati
45 W. Corry Blvd, Edwards 1, Cincinnati OH 45221

**Colin T. Annand (annandct@mail.uc.edu)**
Department of Psychology & Center for Cognition, Action and Perception, University of Cincinnati
45 W. Corry Blvd, Edwards 1, Cincinnati OH 45221



## Abstract

Dotov, Nie and Chemero (2010) conducted a set of experiments to demonstrate how phenomenology, particularly the work of Martin Heidegger, interfaces with experimental research in embodied cognitive science. Specifically, they drew a parallel between Heidegger's notion of readiness-to-hand and the concept of an extended cognitive system (Clark 2008) by looking for the presence or absence of interaction-dominant dynamics (Holden, van Orden, and Turvey 2009; Ihlen and Vereijken 2010) in a hand/mouse system. We share Dotov, Nie and Chemero's optimism about the potential for cross-pollination between phenomenology and cognitive science, but we think that it can be better advanced through a shift in focus. First, we argue in favor of using Maurice Merleau-Ponty's phenomenological theory as the philosophical foundation for experimental research in embodied cognitive science. Second, we describe an audio-visual tracking task in virtual reality that we designed and used to empirically investigate human-environment coupling and interactivity. In addition to providing further support for phenomenologically-inspired empirical cognitive science, our research also offers a more generalizable scientific treatment of the interaction between humans and their environments.



**Keywords:** phenomenology; embodiment; interactivity; agent-environment systems


## Introduction

Dotov, Nie and Chemero (2010) illustrated how insights from the philosophical tradition of phenomenology can contribute to experimental research in embodied cognitive science. In a set of experiments, they had participants play a computer game using the mouse cursor to herd a moving target to a designated area of the screen. In the middle of the experiment, the connection between cursor and mouse was briefly "broken," making the cursor move randomly on screen, independently of mouse movements of the participant, until, after a short period of time, normal operation was resumed. The authors recorded time series data of the mouse/hand position and subsequently submitted it to a detrended fluctuation analysis (Kantelhardt et al., 2001), which estimates a measure of temporal correlation within a time domain signal. They found that when the mouse malfunctioned, there was a shift in the fractal scaling of the mouse/hand movements which they took to correspond to the degeneration of interaction-dominant dynamics into component-dominant dynamics (Holden, van Orden, and Turvey 2009). Following Heideggerian phenomenology, they framed this as a transition from the mouse being *ready-to-hand* to being *present-at-hand* for the participant.

For Heidegger (1927), we perceive objects and tools pragmatically as "something in-order-to": for example, you experience the sheet of paper on your desk as *something to write on* and the pen as *something to write with*. In typical circumstances, these objects are "ready-to-hand" in that, while using them, you can focus on the end goal (writing a letter, say) without having to explicitly attend to the tools themselves. But if something goes wrong and the pen runs out of ink, for example, then the pen becomes "present-at-hand": i.e., it suddenly comes to the forefront of your attention, as something that needs to be confronted explicitly and directly before you can resume your work. It is in this sense that Dotov, Nie and Chemero characterize the mouse in their experiment as shifting from being "ready-to-hand" to "present-at-hand" when it becomes unresponsive.

We agree with Dotov, Nie and Chemero about the potential for cross-pollination between phenomenology and cognitive science. In this paper we explore how this interdisciplinary collaboration can be further promoted through a shift in focus. The object of our investigation is perception-action coupling and interactivity in agent-environment systems: as we propose, understanding how agents engage with features of their environment encompasses a broader range of cognitive phenomena that includes, but is not limited to, tool use. In what follows, we first present Maurice Merleau-Ponty's phenomenological theory as providing the philosophical foundation for this shift. Next, to illustrate what this shift looks like experimentally, we describe findings from a novel audio-visual tracking task in virtual reality that we created. We conclude by discussing how this approach offers a more widely-applicable perspective for phenomenologically-inspired empirical research in embodied cognitive science.

## Merleau-Ponty, Embodiment and Interactivity

Maurice Merleau-Ponty (1945) introduces the term "bodily schema" to describe the "sensori-motor unity of the body" (p. 114). This unity entails, at once, the integration of each of our senses with one another and the integration of perception with action. Seeing and hearing are "pregnant with one another" and they work together as much as our two eyes complement one another. At the same time, seeing and hearing operate in conjunction with our legs and arms to produce walking and grasping: "my body is, not a collection of adjacent organs, but a synergic system, all the functions of which are exercised and linked together in the general action of being in the world" (p. 272).

Merleau-Ponty's famous example of the blind person navigating the environment with a cane or stick shows how this integrated bodily schema is *fluid* and can *change* over time. If you are adept at getting around using a stick, that is because you no longer perceive the stick itself but you perceive the world "at the end of the stick," which involves an expansion of your integrated sensorimotor bodily schema:

> "To get used to a hat, a car or a stick is to be transplanted into them, or conversely, to incorporate them into the bulk of our own body. Habit expresses our power of dilating our being-in-the-world, or changing our existence by appropriating fresh instruments" (1945, p. 166).

Considered by itself, Merleau-Ponty's example of the blind person's cane is compatible with Heidegger's ideas reviewed above: after all, the cane could be said to be ready-to-hand to the expert user in normal circumstances whereas it would become present-at-hand if it suddenly broke in half, just as it would also be initially present-at-hand to a sighted adult who was trying out the cane for the first time while blindfolded. Yet, Merleau-Ponty's understanding of the bodily schema is much broader than the blind man's cane example suggests and, for this reason, it is also better suited for informing empirical research in embodied cognitive science.

First, although the Heideggerian notions of readiness-to-hand and presence-at-hand help make sense of how we use tools (as in the case studied by Dotov, Nie and Chemero), it is not at all clear how this understanding generalizes to a broader range of cognitive phenomena, such as ordinary instances of perception and action that do *not* involve tool use. In contrast, Merleau-Ponty's richer notion of bodily schema is more versatile, applying to embodied experience no matter the degree of "dilation" and regardless of whether it involves the incorporation of tools. In a telling passage, Merleau-Ponty claims:

> "In the gaze we have at our disposal a natural instrument analogous to the blind man's stick. The gaze gets more or less from things according to the way in which it questions them, ranges over or dwells on them." (1945, p. 177).

As this quote suggests, Merleau-Ponty sees our body and our senses as being tool-like in their instrumental or functional character; yet the bodily schema explicitly applies primarily to our basic embodied activity and only secondarily to literal tool use (such as using a hammer or a mouse) as a particular type of bodily activity.

Second, besides applying to a broader range of cognitive phenomena, Merleau-Ponty's notion of bodily schema is also more theoretically attractive because of how it relates to different views in ongoing debates in cognitive science. Dotov, Nie and Chemero interpreted the ready-to-hand mouse as forming, with the body, an *extended cognitive system*. With this, they explicitly tied their account to the hypothesis that cognition may *sometimes* "leak out" of an individual and into parts of the world that the individual is interacting with (Clark 2008). The extended cognition hypothesis is contentious, to say the least: for many cognitive scientists, cognition just is the name of the processing that goes on within the individual's mind/brain; and for advocates of radical embodied cognitive science (e.g., Chemero 2009), the proper object of study just is the animal-environment system as a whole (Gibson 1979).

Merleau-Ponty's notion of bodily schema does not entail a commitment to the contentious hypothesis of extended cognition, and it thereby circumvents the controversy. In a key passage, Merleau-Ponty explains: "With the notion of the bodily schema we find that not only is the unity of the body described in a new way, but also, through this, the unity of the senses and of the object" (1945, p. 273). Above we saw that the bodily schema entails the sensorimotor unity of the body, that is, the integration of the senses and between perception and action. This quote adds, further, that the bodily schema entails also an integration between subject and the objects of experience. This captures an essential feature of the radical embodied and Gibsonian approaches to studying agent-environment systems, namely the focus on the complex interactivity between agent and environment: in this view, "patterns of an organism's behavior are best understood as the emergent property of the interactions of the organism with its environment" (Kelty-Stephen, Palatinus, Saltzman, and Dixon 2013, p. 2) and "perception and action are best understood in the broader context of the task and environment within which coordination of those biological nuts and bolts takes place" (p. 3). As such, we suggest, embodied agency or "being in the world" is always characterized by an integration of agent and environment through interaction. Interactivity may change qualitatively with changes in task and in the availability of task-relevant information, but it is always present: an agent's perception-action never becomes fully detached from her environment, and understanding this relation is independent of whether some internal feature of the agent "leaks out" into the world or not.

As an illustration of interactivity, imagine an ordinary situation such as trying to track a bumble bee so as to avoid being stung. Although you may initially catch sight of the bee and follow it with your gaze, the bee's erratic movement might cause it to disappear against a cluttered background.

Your desire to avoid being stung persists and you maintain an awareness of the bee's position by listening, trying to regain sight of it. You swivel your head, accommodating for the subtle shifts in interaural sound intensity, allowing your ears to guide your continued search for the bee. Furthermore, the bee may fly along your sagittal plane, momentarily escaping your efforts to track it by sound until, finally, you are able to regain auditory or visual tracking. This dance, between you and the bee, may persist until the bee exits your immediate surroundings. Although it may be true that, at times, the differences in mode and strength of your sensory coupling to the bee change, nothing is ever "broken." There may be differences in how your head or eyes move relative to the bee, but no aspect of this system can be said to transition from readiness-to-hand to presentness-at-hand. Furthermore, the system maintains interactivity throughout. Even though the specific dynamics of a particular aspect of the system may change, the system continues to be unified through the ongoing pursuit of the goals that are implicit to the task (e.g. avoiding a sting). The experiment described below was designed to capture this point.

## Method

Undergraduate students (N = 10) at the University of Cincinnati participated in a virtual audio-visual tracking task for class credit. At the start of the experiment, each participant put on an Acer mixed reality headset and a pair of in-ear monitors (IEMs). The virtual scene that they were presented with (depicted in Figure 1) consisted of a white room and a semi-circular line spanning 180 degrees of the participant's visual field on which a black and yellow sphere (henceforth, "the bee") would travel over time, moving in a roughly brownian fashion similar to the moving target from Dotov, Nie and Chemero (2010).

Participants were instructed to track the bee throughout the task by continuously pointing their center of vision, indicated by a small sphere, to its location. They were told that the source would begin emitting a buzzing sound when the experiment started and that it would be necessary to use this sound to continue tracking the bee because the bee would shortly become invisible. The spatial information present in the sound of the bee was imparted by a set of generalized head-related transfer functions (Zhong and Xie 2014). Unbeknownst to the participants, after the bee had been invisible for 12 seconds, the sound spatialization would be removed, making it impossible for the participant to effectively track the bee. After a period of time, the sound spatialization would be added back and then, finally, the bee would reappear. In total, the task consisted of two 12 second periods of audio-visual tracking (at the beginning and at the end), two 12 second periods of audio-only tracking, and one 12 second period of tracking with no spatial information (in the middle). The order of this sequence is illustrated in Figure 2. During the entire experiment, the angular difference between the participant's center of vision and the position of the bee was recorded at 100 hz.

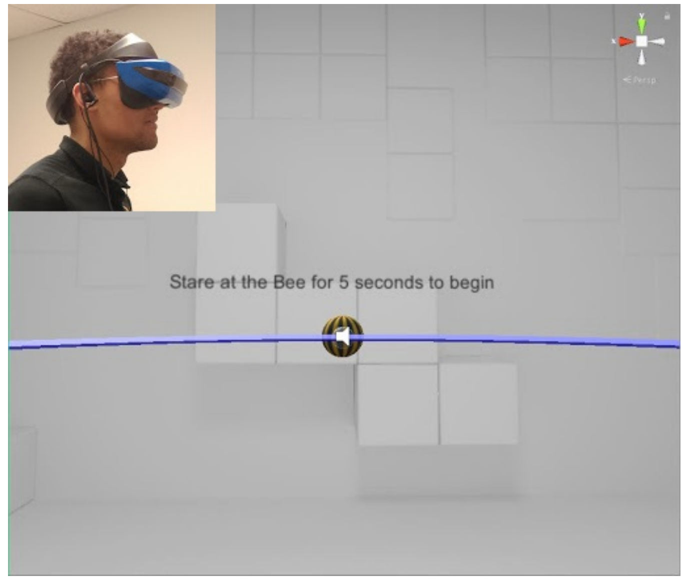


Figure 1: An image of what a participant would see upon starting the experiment, including the bee, the instructions for starting a trial and the line upon which the bee moved.

### Fractal Analysis

We submitted the time series data from each trial and condition (Audio-Visual Information (AV), Audio Information Only (AO), and No Spatial Information (NI)) to a detrended fluctuation analysis (DFA) which allows for temporal correlations within a signal, at different scales, to be captured by a single value. Detrended Fluctuation Analysis (DFA) is a form of fractal analysis, which describes a power-law that captures the relationship between the size and occurrence rate of fluctuations for a given time series (Ihlen 2012). Fractal analyses have previously been used to illuminate the nature of embodied cognitive activity by examining continuous measures of agents embedded in environments (Kello, Beltz, Holden & Van Orden 2007).

The DFA measurement of the time series of angular error between gaze and target for each information condition (AV, AO and NI) yields a Hurst exponent and a closely related Alpha value, both of which describe the power-law relationship within the time series. In Dotov, Nie and Chemero (2010), Alpha values were calculated at repeated intervals to identify changes in tool-use behavior that were caused by the perturbation of mouse function. Here, we calculated Hurst exponents for each condition in order to index how gaze activity changes across the information conditions in the bee tracking task, as is visually exemplified in the time series data shown in Figure 2. Our DFA used a minimum window size of 2 samples and a maximum window size of roughly one third of each condition time series, which were each 1200 samples in length.

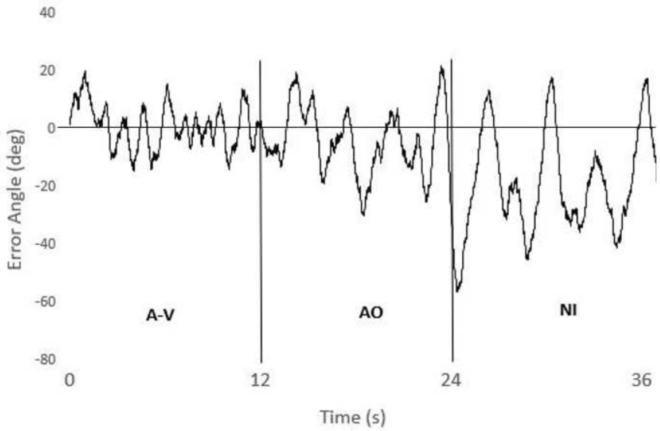


Figure 2: The timeseries above shows the first 36 seconds (x-axis) for the error angle (y-axis) between a participant's center of vision and the bee during the tracking task. The different information conditions are indicated by the lines, Audio and Visual information (AV), Audio information only (AO) and No spatial information (NI). In this particular case, the size and frequency of error clearly changes between the information conditions: error is minimized in the audio-visual condition (AV), increases in the auditory only condition (AO), and displays large, shifting values in the no spatial condition (NI) when the bee target auditory signal switches from stereo to mono.

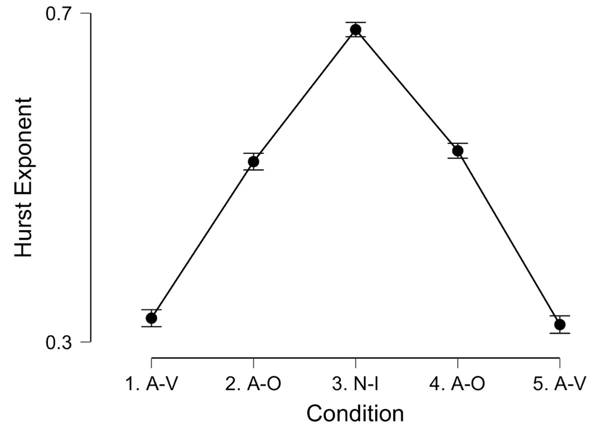


Figure 3: The x-axis indicates the changing conditions between audio-visual information (AV), auditory information only (AO), and no spatial information (NI). The y-axis indicates the value of the mean Hurst exponents and standard error bars for each condition.

## Results

The Hurst exponents, calculated for each information condition and trial (AV, AO and NI) were submitted as dependent variables to a repeated measures analysis of variance to examine changes across trials as well as conditional differences.

Neither the within subject main effect of Trial, $F(4,180) = 0.125$, $p = 0.97$, or Trial by Condition interaction effect, $F(16,180)$, $p = 0.1$ were significant. The effect of Condition was significant, $F(4,45) = 86.38$, $p < .001$, $n^2 = 0.88$. It is worth noting that Tukey post-hoc analysis revealed significant differences between the condition types, but not their separate time occurrences: both AV conditions are not significantly different from one another, and both AO conditions are not significantly different from one another. Further details are provided in Table 1 and in Figure 3.

Table 1: Descriptive Statistics

| | Mean Hurst Exponents by Condition | | | | |
|---|---|---|---|---|---|
| | 1. A-V | 2. A-O | 3. N-I | 4. A-O | 5. A-V |
| Mean | 0.3290 | 0.5194 | 0.6800 | 0.5326 | 0.3212 |
| Std. Dev. | 0.0641 | 0.0406 | 0.0327 | 0.0459 | 0.0664 |

## Discussion

Our experimental results reveal differences in the fractal scaling of movement across shifting task conditions. In this way, our results were similar to what was found by Dotov, Nie and Chemero (2010). The key difference between the two experiments lies in our focus. Their investigation is centered on the agent; ours follows radical embodied cognitive science (Chemero 2009) by being primarily concerned with the agent-environment system as a whole. This difference in focus informed both our choice of dependent measure and our interpretation of interactivity.

Because they were trying to find support for the extended cognition hypothesis, Dotov, Nie and Chemero (2010) measured raw hand movement at the tool/hand interface. This meshes well with their goal of demonstrating a shift, from the agent's perspective, between a tool being ready-to-hand to becoming present-at-hand—but this approach misses out on capturing the rest of the agent-environment system. In contrast, we adopted a collective measure at the task performance level. By measuring the error angle between the gaze and the bee's position, we were able to detect shifts in the overall agent-environment dynamics. In this context, specific Hurst exponent values are useful and explicate the nature of the system. For example, in the Audio-Visual Information condition, the low Hurst value indicates that the system corrects for increases in error similarly across timescales, exhibiting anti-persistent dynamics (Riley et al 2012). This makes sense because participants are likely very good at visually orienting to the position of objects. The higher Hurst values from the Auditory Only and No Information conditions show that there is a shift in how error is accommodated for at different scales. In the Auditory Only condition, for example, the participant may be able only to accommodate for movements of the bee very slowly, but is ineffective at tracking its faster movements. This shift can be characterized as a shift towards persistent system dynamics (Riley et al 2012), which continues in the same direction as information is reduced further in the No Information condition.

A similar interpretation could have been applied to the herding task of Dotov, Nie and Chemero (2010) if the dependent measure had reflected the collective dynamics of the agent-environment system. In their case, it's not that when the tool breaks it is noticed as a tool, external to the system. Rather, the tool appears broken within the context of a task and is used as such. Movements exhibited by participants experiencing a broken mouse are sensible as movements meant to fix or disambiguate the nature of the brokenness of the mouse. Similarly, the movements of our participants who had no information about the bee's position are sensible as exploratory procedures (Riley et al 2002), i.e., movements meant to pick up information. These movements do not reflect a degeneration of interaction, but only a shift in the nature of the ongoing interaction between agent and environment. A participant in either task is never truly decoupled from the specific environment implied by the overarching task.

Dotov, Nie and Chemero characterize the distinction between interaction-dominant dynamics and component-dominant dynamics as follows: "In component-dominant dynamics, behavior is the product of a rigidly delineated architecture of modules, each with predetermined functions; in interaction-dominant dynamics, on the other hand, coordinated processes alter one another's dynamics, with complex interactions extending to the body's periphery and, sometimes, beyond" (2010, p. 3). This characterization works well with their agent-centered approach and their focus on cognition as an internal feature of the agent that can potentially extend out into the world. But when the object of study becomes the agent-environment system, as proposed in radical embodied cognitive science (Chemero 2009), this characterization fails. The dynamic variation in a proper collective measure of a complex agent-environment system will always be governed by the interaction between agent and environment. The system may be redefined across tasks, but can never become broken in the way that Dotov, Nie and Chemero would require. Because interactivity is a universal feature of agent-environment systems, rather than looking for signs of a shift from interaction-dominance to component-dominance, it is more appropriate to inquire into the specific nature of the interactivity. This means focusing on task specific coordination (Turvey, Saltzman and Schmidt 1991), rather than the dynamics that play out at the interface between human and tool.

As seen above, the choice of focus of investigation—whether centered on the agent or on the agent-environment system as a whole—is directly linked to the choice of dependent measure and to the interpretation of interactivity. The focus of investigation is also intimately associated to the phenomenological theory adopted in each case. Heidegger's theory is agent-centric and lends itself to application for investigating the dynamics of tool use and cognitive extension. Merleau-Ponty's theory, on the other hand, motivates thinking in terms of an integration between subject and object, or between agent and environment. This makes it more apt for making sense of a broader range of cognitive phenomena, beyond tool use, where interaction may occur. Merleau-Ponty's approach is thus better suited for conceptually framing research into perceptually driven human-environment interactivity in the ecological and embodied cognitive sciences.

## References


Chemero, A. (2009). *Radical embodied cognitive science*. MIT press.

Clark, A. (2008). *Supersizing the mind: Embodiment, action, and cognitive extension*. Oxford University Press.

Dotov, D. G., Nie, L., & Chemero, A. (2010). A demonstration of the transition from ready-to-hand to unready-to-hand. *PLoS One*, *5*(3), e9433.

Gibson, J. J. (1979). *The ecological approach to visual perception*. Houghton Mifflin.

Heidegger, M. (1927/2001). *Being and time* (J. Macquarrie & E. Robins, Eds.). Blackwell Publishers Ltd.

Holden, J. G., Van Orden, G. C., & Turvey, M. T. (2009). Dispersion of response times reveals cognitive dynamics. *Psychological review*, *116*(2), 318.

Ihlen, E. A., & Vereijken, B. (2010). Interaction-dominant dynamics in human cognition: Beyond 1/α fluctuation. *Journal of Experimental Psychology: General*, *139*(3), 436.

Kantelhardt, J. W., Koscielny-Bunde, E., Rego, H. H., Havlin, S., & Bunde, A. (2001). Detecting long-range correlations with detrended fluctuation analysis. *Physica A: Statistical Mechanics and its Applications*, *295*(3-4), 441–454.

Kello, C. T., Beltz, B. C., Holden, J. G., & Van Orden, G. C. (2007). The emergent coordination of cognitive function. *Journal of Experimental Psychology: General*, *136*(4), 551568.

Kelty-Stephen, D. G., Palatinus, K., Saltzman, E., & Dixon, J. A. (2013). A tutorial on multifractality, cascades, and interactivity for empirical time series in ecological science. *Ecological Psychology*, *25*(1), 1-62.

Merleau-Ponty, M. (1945). *Phenomenology of perception* (. Translation by Kegan Paul, Ed.). Routledge.

Riley, M. A., Bonnette, S., Kuznetsov, N., Wallot, S., & Gao, J. (2012). A tutorial introduction to adaptive fractal analysis. *Frontiers in physiology*, *3*, 371.

Riley, M. A., Wagman, J. B., Santana, M. V., Carello, C., & Turvey, M. T. (2002). Perceptual behavior: Recurrence analysis of a haptic exploratory procedure. *Perception*.

Turvey, M. T., Saltzman, E., & Schmidt, R. C. (1991). Dynamics and task-specific coordinations. In *Making them* *move: Mechanics, control, and animation of articulated figures* (p. 157-170).

Van Orden, G. C., Holden, J. G., & Turvey, M. T. (2005). Human cognition and 1/f scaling. *Journal of Experimental Psychology: General*, *134*(1).

Zhong, X.-l., & Xie, B.-s. (2014). Head-related transfer functions and virtual auditory display. In *Soundscape semiotics-localization and categorization.* IntechOpen.