

Prediction and experimental protocols in neuroscience: a case study

Edoardo Datteri
Dipartimento di Scienze Umane “R. Massa”,
Università degli Studi di Milano-Bicocca
e-mail: edoardo.datteri@unimib.it

1. Introduction
2. The domain of validity of neuroscientific generalizations
3. Why place cells fire where they fire
4. Discussion

ABSTRACT. Neuroscientific generalizations are highly idealized, in the sense that they do not explicitly mention the large set of factors that may perturb the stated regularities. This gives rise to the conceptual problem – which has been often addressed and discussed in the philosophical literature on models – of understanding how they can be tested by experiments performed in non-ideal conditions. This paper emphasizes the role of experimental protocols in the testing of idealized neuroscientific generalizations by appeal to “real-world” experiments, in connection with a case study on the behaviour of rat place cells.

1. Introduction

Many neuroscientific investigations on the behaviours and the cognitive capacities of living systems involve the formulation and the experimental testing of generalizations: the long-standing research trend on the mechanisms underlying spatial memory in rats is a case in point. It is well known that the so-called place cells in area CA1 of the rat hippocampus “fire only when an animal moves through a particular location in space” (Moser and Moser 2008, 1142). Some years after the discovery of place cells, head-direction (HD) cells were discovered in the rat brain: “Each [HD] cell fires whenever the rat faces one particular direction, regardless of where the rat is within the environment or what it is doing. Different cells have different preferred directions, so that,

for each possible directional heading, there is a unique set of HD cells that will be active” (Sharp et al 2001, 289). In 2004, another striking regularity in the relationship between rat position and brain cell activity has been found to hold in the medial entorhinal cortex (MEC), an area that is anatomically connected to CA1 and to other areas in the hippocampus: “A key cell type of the MEC is the grid cell, the spatial receptive fields form a regularly tassellating triangular pattern that spans the complete environment covered by the animal” (Fyhn et al 2007, 190): each grid cell fires whenever the rat is on the vicinity of the vertex of a triangular grid which covers the whole environment.

These generalizations are highly idealized, in the sense that they do not explicitly mention the (large) set of factors that may perturb the regularity stated there. This gives rise to the conceptual problem – which has been often addressed and discussed in the philosophical literature on scientific models – of understanding how they can be tested by experiments performed in *non-ideal conditions*. This paper emphasizes the role of *experimental protocols* in the testing of idealized neuroscientific generalizations by appeal to “real-world” experiments, in connection with a case study on the behaviour of rat place cells (O’Keefe and Conway 1978).

2. The domain of validity of neuroscientific generalizations

Consider the following *non-probabilistic* generalization, which literally posits a *strict* relationship between the activity of place cells and the position of the animal in space.

(G) Each place cell fires only if the rat is located in a particular spatial position.

This generalization includes no explicit *ceteris paribus* clause.¹ One may well suspect that (G) is elliptical as it stands, as it does not state explicitly the (large) set of perturbing factors that must be idle for place cells to fire only

¹ As discussed by many authors (see Earman et al 2002), one may reasonably worry that generalizations explicitly including a non-eliminable *ceteris paribus* clause are in principle not testable, as the clause makes the statement vacuously true. Similar concerns may arise when a regularity is explicitly asserted to hold “typically”, “normally”, and the like. *Ceteris paribus* generalizations are not particularly common in neuroscience; the present discussion is restricted to literally strict generalizations like (G).

when the rat is in the corresponding receptive field. A way to express this claim is to assert that generalization (G) states a correlation between place cells activity and rat position which obtains only in *ideal settings*, i.e., in situations in which no relevant perturbing factor is at work.

This generalization gives rise to a *prima facie* testability problem (Earman et al 2002; Pietroski and Rey 1995; Giere 1999; Suppe 1984; Redhead 1980). It is not clear why experimental counter-examples – i.e., place cells firing when the rat is not in the corresponding position – should be taken to count against (G), as they could be easily explained away by assuming that they are due to perturbing conditions that are unmentioned in (G). More generally, it is not clear why a counter-example obtained in non-ideal settings should be brought to count against a generalization that states how things would go in the ideal setting in which no perturbing condition is at work.

This *prima facie* testability problem is really a problem if one has no idea at all, or only a vague or imprecise idea, of what the relevant perturbing factors are. According to Woodward (2000) the latter is indeed the case in the special sciences: often “scientists can give a simple, precise, and general characterization of the domain in which a generalization holds ... But more commonly, especially in the special sciences, such a general characterization will be unknown and may not even exist at the level at which one is theorizing” (231). In many cases, the circumstances in which a generalization is likely to break down “are specified in an informal and rather imprecise way”; the restrictions on the validity of a generalization are often “vague and imprecise – they are best viewed as rules of thumb rather than as specifications of the exact circumstances in which we should expect [the generalization] to hold” (233). One may well expect this vagueness to determine some uncertainty in the interpretation of apparent counter-examples: what criteria one may have to decide whether an experimental counter-example has been obtained within the domain of validity of the target generalization (which would make it a *proper* counter-example) or is due to the violation of some relevant boundary condition (which would make it a poor basis to test the idealized generalization)?

The study carried out by O’Keefe and Conway (1978) provides insights to reflect on this issue. The authors do not provide a set of precise and testable statements expressing the nature and the magnitude of the possible perturbing factors (Woodward’s claims on the fuzziness of the domain of validity of special science generalizations applies quite well there); however, they seem to have a perfectly clear idea – something much stronger than “very rough rules of thumb” – on how to interpret even *one* apparent prediction failure on the

behaviour of *one* neuron, and on how to justify their theoretical conclusions accordingly. Let's see.

2. Why place cells fire where they fire

The main goal of the O'Keefe and Conway's (1978) article is to understand what types of stimuli or states of affair are regularly correlated with place cells activity. The authors have some working hypotheses, stated in the form of *highly elliptical* generalizations, including the following.

1. Each place cell is active whenever the rat is in a particular position in space with respect to some sensory cues.
2. Each place cell is active whenever the rat is in a particular position with respect to the overall structure of the maze (independently of any sensory cue).

Hypotheses (1) and (2) are tested by recording CA1 activity of rats running freely in a T-shaped maze, located inside an experimental enclosure delimited by curtains. There are three sources of sensory stimuli inside, each one mounted on a different wall of the enclosure: a lamp, a fan, a buzzer, and a card. The start arm can be placed on either side of the longest segment of the T (see Figure 1).

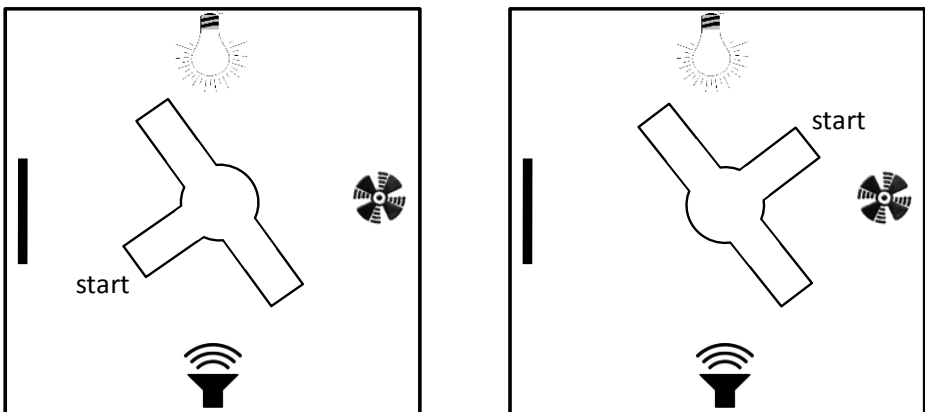


Figure 1

Notably, a *finite* set of *explicit*, *well-defined* and *testable* procedural rules is chosen to define a “proper” experimental setting. One of these rules is the following (see the target article for other examples):

The whole enclosure – together with the T-maze inside – must be rotated from trial to trial with respect to the external environment to avoid any possible influence of geomagnetic fields (see Figure 2).

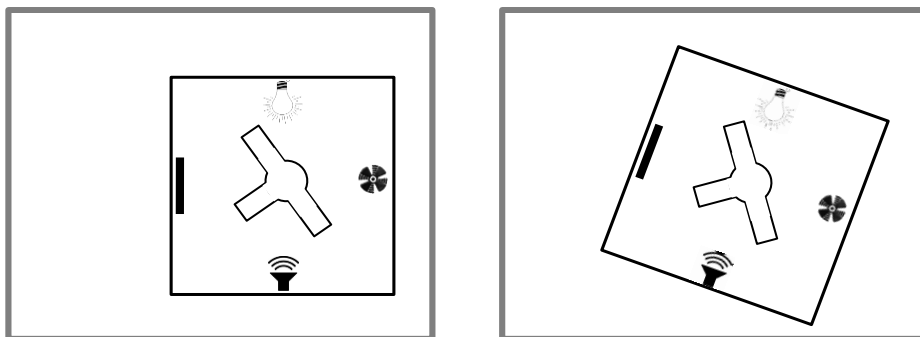


Figure 2

Generalization (1) is supported by many experimental recordings, which also contribute to rejecting generalization (2): place cells firing in the start arm stop firing when the start arm is placed on the other side of the longest segment (this suggests that place cells firing depends on the rat’s position *with respect to the sensory cues*, as generalization 1 prescribes, contrary to generalization 2, according to which those place cells should fire when the rat is in the “start arm” independently of its position with respect to the cues). Notably, however, unit #23 is found to apparently contradict these results: it fires in the start arm on either side of the cross bar, consistently with generalization 2. Is this result to be regarded as a proper counter-example to hypothesis 1? The authors conclude in the negative: they discard this result and explicitly justify this choice by appeal to reasons which have nothing to do with the mere desire to save hypothesis 1. These reasons are connected to the fact that one of the procedural rules we have just mentioned – rotating the enclosure to avoid possible effects of geomagnetism – has not been properly followed when recording the behaviour of unit #23 (see discussion on pp. 584-585 of the target article). So, one cannot exclude that this result is due to some (in this case, geomagnetic) perturbing condition. In other experimental sessions,

the behaviour of another cell – unit #25 – is found to contradict another previously hypothesized generalization (i.e., that place cells stop firing when all the sensory cues are removed). All the pre-defined procedural rules have been properly applied in this case (see p. 586): accordingly, this experimental result is not discarded but taken as a proper counter-example to that generalization.

2. Discussion

It is unquestionably true that neuroscientific generalizations – such as those involved in the selected study – are highly elliptical: their domain of validity is not defined in terms of a set of well-defined boundary conditions (e.g. “generalization 1 holds provided that the level of geomagnetism is between ... and ...”). This is not to say, however, that their domain of validity is not defined, nor that neuroscientists have at most “very rough rules of thumb” (Woodward 2000, 231) to assess whether they are within that domain or not: they choose an experimental protocol made up of reasonably well-defined and testable rules, and these rules define the domain of validity of their generalizations in a procedural way. In other words, an elliptical (idealized) generalization can be evaluated on the basis of a “real-world” experimental result provided that the procedural rules defining the domain of validity of the generalization have been properly followed. If the latter is the case, counter-examples can be taken to count against the target generalization; otherwise, results can be discarded as irrelevant to draw a theoretical conclusion on the target generalization. The testable character of the procedural rules – which are more than just “rough rules of thumb” – provides one with good (not ad hoc) reasons to decide on whether “real-world” experimental results can be brought to bear on generalizations stating how things would be in ideal circumstances, i.e., in situations in which no relevant perturbing factor is at work.²

RIFERIMENTI BIBLIOGRAFICI

EARMAN, J., ROBERTS, J. T., & SMITH, S. (2002): “Ceteris Paribus Lost”, *Erkenntnis*, 57(3), p. 281–301.

² Experimental protocols play, at least in this study, (part of) the role ascribed by Suppe (1989) to what he calls the *theory of experimental design* in the application of idealized theories to the world.

PREDICTION AND EXPERIMENTAL PROTOCOLS IN NEUROSCIENCE:
A CASE STUDY

- FYHN, M., HAFTING, T., TREVES, A., MOSER, M. B., & MOSER, E. I. (2007): "Hippocampal remapping and grid realignment in entorhinal cortex", *Nature*, 446(7132), p. 190.
- GIERE, R. N. (1999): *Science without laws*, Chicago: University of Chicago Press.
- MOSER, E. I., & MOSER, M.-B. (2008): "A metric for space", *Hippocampus*, 18(12), p. 1142-56.
- O'KEEFE, J., CONWAY, D. H. (1978): "Hippocampal place units in the freely moving rat: Why they fire where they fire", *Experimental Brain Research*, 31(4), p. 573-590.
- PIETROSKI, P., & REY, G. (1995): "When other things aren't equal: saving ceteris paribus laws from vacuity", *The British Journal for the Philosophy of Science*, 46(1), p. 81.
- REDHEAD, M. (1980): "Models in physics", *The British Journal for the Philosophy of Science*, 31(2), p. 145.
- SHARP, P. E., BLAIR, H. T., & CHO, J. (2001): "The anatomical and computational basis of the rat head-direction cell signal", *Trends in Neurosciences*, 24(5), p. 289-294.
- SUPPE, F. (1989): *The semantic conception of theories and scientific realism*, Champaign: University of Illinois Press.
- WOODWARD, J. (2000): "Explanation and Invariance in the Special Sciences", *British Journal for the Philosophy of Science*, 51, p. 197-254.
- WOODWARD, J. (2003): *Making things happen: A theory of causal explanation*, New York: Oxford University Press.