MAKE INFORMATION IN SCIENCE MEANINGFUL AGAIN

Javier ANTA

ABSTRACT: Although the everyday notion of information has clear semantic properties, the all-pervasive technical concept of Shannon information was defended being a non-semantic concept. In this paper I will show how this measure of information was implicitly 'semantized' in the early 1950s by many authors, such as Rothstein's or Brillouin's, in order to explain the knowledge dynamics underlying certain scientific practices such as measurement. On the other hand, I will argue that the main attempts in the literature to develop a quantitative measure of semantic information to clarify science and scientific measurements, such as Carnap-Bar-Hillel, or Dretske, will not successfully achieve this philosophical aim for several reasons. Finally, I will defend the use of a qualitative notion of semantic information within the information-theoretical framework MacKay to assess the informational dynamics underlying scientific practices, particularly measurements in statistical mechanics.

KEYWORDS: Shannon information, communication theory, measurement, semantic information

1. Introduction

It is often said that the task of a scientist is to obtain information about his or her field of research. But what do we mean here by 'information'? As the information philosopher Floridi[1] argues, the term 'information' encompasses a huge plurality of different concepts and different meanings.[2] One of these is the ordinary sense of information, of an indisputable semantic character (highlighted in his 'aboutness,' i.e., John has information *about* what happened) and linked to the ability to provide knowledge to someone. However, Claude Shannon's famous theory of communication[3] placed his quantitative concept of information at the center of the imaginary of the technical-scientific community since the 1950s, whose importance will expand until today. Interestingly, this author employed a measure-concept of

[1] See Luciano Floridi, *The Philosophy of Information* (Oxford University Press, 2011)

[2] See Olimpia Lombardi, Federico Holik, and Leonardo Vanni, "What is Shannon information?" *Synthese* 193, 7 (2016):1983-2012.

[3] See Claude E. Shannon, "A Mathematical Theory of Communication," in *Collected Papers*, eds. N. J. A. Sloane and A. D. Wyner (New York: IEEE Press, 1948 [1993]).

information from a certain sequence of symbols that was completely independent of the meaning of these symbols, which contradicted the semantic character of the ordinary meaning of this term.

The enormous intellectual impact that Shannon's theoretical proposal had prompted several authors to use the concept of theoretical-communicative information to understand how scientists acquire information about their objective phenomena. For example, Jerome Rothstein in 1951 systematically compared scientific measurement processes with a communication process between observer and observed system.[4] However, as I will argue in this paper, this kind of application of Shannon information would implicitly involve attempts to semantize (in a misleading fashion) the everyday meaning of this term. In this paper I will argue that the main attempts in the literature to employ a theoretical-communicative concept of information (or other alternatives also statistical-quantitative) to illuminate scientific practices will not succeed in carrying out this task in a robust way, even if conceptual alternatives to Carnap and Bar-Hillel[5] are developed or formally complemented to Dretske[6] in a way that is somehow sensitive to the semantic content of certain information elements. On the other hand, I will propose not a new quantitative measure of information adapted to these scientific scenarios, but rather to develop an informational interpretation (using certain elements developed by MacKay[7]) of the very representational tools used in scientific practices.

The plan for this paper is the following. Next, I will present the basis of Shannon's communication theory,[8] where a non-semantic concept of information was developed in order to statistically evaluate the transmission of signals. In Section 3, I will analyze how, despite these characteristics, Shannon's informational concept began to be used during the 1950s as an informational measure used to clarify certain scientific practices such as measurement,[9] implicitly providing it with a semantic character. Later on, I will detail the main quantitative-statistical theories of information developed under the intellectual impact (either as an alternative or as a complement) of Shannon's proposal. In Section 5 I will argue that this first and last would fail to give a satisfactory account for various reasons of the informational

---

[4] Jerome Rothstein, "Information, Measurement, and Quantum Mechanics," *Science* 114 (1951): 171-175.

[5] Rudolf Carnap and Yehoshua Bar-Hillel, "An Outline of a Theory of Semantic Information," *Technical Report* 247 (1952), Research Laboratory of Electronics, MIT.

[6] Fred Dretske, *Knowledge and the Flow of Information* (Cambridge: MIT Press, 1981).

[7] Donald MacKay, *Information, Mechanism, and Meaning* (Cambridge: MIT Press, 1969).

[8] Shannon, "A Mathematical Theory."

[9] See Rothstein, "Information, Measurement" and Leon Brillouin, *Science and Information Theory* (New York: Academic Press, 1956).

dynamics of scientific practice. Finally, I will develop an alternative informational proposal from MacKay's framework based on the definition of a qualitative concept of information, in order to philosophically assess certain scientific measurement practices.

2. Shannon's Communication Theory

In "A Mathematical Theory of Information," Claude Shannon first set out the foundations of his theoretical proposal regarding the statistical analysis of the transmission of continuous or discrete messages (i.e., sequences of symbols belonging to a set of symbols or alphabet) within certain communicative contexts.[10] These contexts are made up of (i) a 'source' that generates the message, (ii) a 'transmitter' that transforms the message into signals to be transmitted, (iii) a 'receiver' that reconstructs the message at the point of destination, and (iv) a 'communicative channel' as a means of transmitting messages. In not so general terms, his proposal is based fundamentally on developing an H function (this is called 'source entropy') that measures the average amount of information generated by the source through the transmission of a message si...sj, where the probability of occurrence of each particular symbol is determined by the probability distribution over the source:

$$\mathrm{H}\ (\mathrm{S}) = -\sum p(s_i) \log p(s_i) \qquad (1)$$

Roughly speaking, this amount of H entropy measures the degree of improbability (also often interpreted in epistemic terms as 'degree of unexpectedness' or 'unpredictability') with which a particular sequence of symbols s_i occurs, reaching its maximum value when the probability distribution over the source is uniform. For example, the message 'XZV' will have a significantly higher entropy than the message 'SKY' if we consider the frequency of occurrence of the English letters (i.e. H(XZV) > H(SKY)), precisely because the simultaneous occurrence of the 'XZV' symbols in a message would be extremely uncommon. Thus, this measure of information cannot be defined only for a particular message (i.e. 'SKY'), but for a message with respect to the probability distribution defined for the source or the set of all possible source symbol sequences. It is precisely the so-called 'noisy coding theorem' that mathematically determines that the most optimal way to encode messages is by sequences of units with binary values (or 'bits') of 0 and 1.

One of the main characteristics of Shannon entropy is its disregard for the semantic content of the message: "Frequently, the messages have meaning; that is they refer to or are correlated according to some with certain physical or conceptual

[10] Shannon, "A Mathematical Theory."

entities. These semantic aspects of communication are irrelevant to the engineering problem."[11] In other words, it is completely irrelevant whether the sequence 'SKY' actually refers to the sky or to a particular brick in the street so that the amount of information measured by means of the Shannon H(SKY) entropy is, for example, 0.05. This lack of semantic sensitivity of Shannon measure[12] of the information of a message to its (plausible) conventional meaning was the subject of enormous controversy since the popularization of this theory in the 1950s until today,[13] mainly because of its radical difference with the usual sense of the term 'information,' synonymous with 'knowledge' and inseparably linked to semantic-epistemic and intentional properties,[14] i.e. information of A *about* X. Another important point to disregard Shannon entropy as a measure of meaning or semantic content of messages (or other structures) would be the following: the function H does not measure the amount of (not semantics, as we have just pointed out) information conveyed in the transmission of particular messages, but the average amount of information of a statistical assembly of possible messages.[15]

Another important aspect of this entropy measure H is its formal similarity (based on the use of probability distributions and the use of the logarithmic function for its pragmatic virtues) with the Boltzmann and Gibbs measure of statistical mechanical entropy "the form of H will be recognized as that of entropy as defined in certain formulations of statistical mechanics."[16] Despite this similarity, the latter measure (roughly) the probability that the exact microscopic state of the physical system is in a cell or region of the space of possible molecular values or phase space. This leads us to consider the choice of this name 'entropy' originally belonging to the field of statistical physics to name a measure of quantity of information. Tribes

---

[11] Shannon, "A Mathematical Theory," 3.

[12] Recently, M. Alistair ("The Semantics Latent in Shannon Information," *British Journal for the Philosophy of Science* 70, 1 (2019): 103-125) has argued that it would be possible to extract certain semantic properties from the statistical correlations modeled from Shannon's theoretical-communicational formalism. However, his argument depends on a proto-information theory developed by Turing in the 1940s, and not directly on Shannon's theory.

[13] See Ronald Kline, *The Cybernetics Moment: Or Why We Call Our Age the Age of Information* (Baltimore: Johns Hopkins University Press, 2015), Section 2.

[14] See Floridi, *Philosophy of Information*.

[15] See Lombardi, Holik and Vanni, "What is Shannon." Note that it is misleading to (although often convenient) say that one or the other message conveys unit information. The concept of information applies not to the individual messages (as the concept of meaning would), but rather to the situation as a whole, the unit information indicating that in this situation one has an amount of freedom of choice, in selecting a message, which it is convenient to regard as a standard or unit amount (Kline, *The Cybernetics Moment*, 132).

[16] Shannon, "A Mathematical Theory," 12.

and McIving reported (from an interview of Shannon) that it was von Neumann who suggested the name to exploit his deep misunderstanding within the scientific community.[17]

As Shannon himself recognized, his proposal is presented as a highly sophisticated development that builds on the analyses developed in the papers by Nyquist and Hartley (where the latter's measure of 'information' is mathematically identical with Shannon's entropy in the case of a uniform probability distribution over the symbols) during the 1920s, deeply forgotten beyond the walls of the Bell Laboratories. Unlike these pioneers, Shannon developed throughout the 1940s (culminating in his 1948 paper) an extensive technical proposal on how to statistically evaluate and optimize the transmission of discrete/continuous messages in both noisy and noiseless channels, modeling this communicative process as a Markov chain. In short, we must emphasize once again that Shannon information is intrinsically independent of the meaning and physical character of the informational elements: "Shannon's theory is a quantitative theory whose elements have no semantic dimension (...) Moreover, Shannon's theory is not tied to a particular physical theory, but is independent of its physical interpretation."[18]

The director of the Division of Natural Sciences at the Rockefeller Foundation, Warren Weaver, immediately appreciated the prospects of Shannon's theory, not only within the field of communication but also in other scientific domains. His role in the enormous immediate impact that communication theory had within the scientific community was pivotal, popularizing Shannon's excessively technical proposal for the general public (note that even engineers had difficulties in understanding his theses[19]) through an introductory commentary in the reprint of the original article in the famous (Shannon and Weaver) 1949 book.[20] Moreover, one of his main aims (more or less implicit) was to elevate Shannon's intellectual work to the Olympus of American science in which the physicist J. W.

---

[17] "[Shannon said:] 'My greatest concern was what to call it. I thought of calling it 'information', but the word was overly used, so I decided to call it 'uncertainty.' When I discussed it with John von Neumann, he had a better idea. Von Neumann told me, 'You should call it entropy, for two reasons. In the first place you uncertainty function has been used in statistical mechanics under that name. In the second place, and more importantly, no one knows what entropy really is, so in a debate you will always have the advantage" (Myron Tribus and Edward McIrvine, "Energy and information," *Scientific American* 225 (1971): 179-188).

[18] See Lombardi, Holik, and Vanni. "What is Shannon," 2000.

[19] Kline, *The Cybernetics Moment*, Chapter 2.

[20] Claude E. Shannon and Warren Weaver, *The Mathematical Theory of Communication* (Urbana: University of Illinois Press, 1949).

Gibbs was, but to generate a historical narrative in which Shannon's theory was the culprit of the physical domain statistical mechanics as developed by Boltzmann:

> Dr. Shannon's work roots back, as von Neumann has pointed out, to Boltzmann's observation, in some of his work on statistical physics (1984) that entropy is related to 'missing information,' inasmuch as it is related to the number of alternatives which remain possible to a physical system after all the macroscopically observable information concerning it has been recorded.[21]

In this direction, Weaver intended to extend the domain of application and popularization of Shannon's theoretical proposal from the narrow technical field of signal transmission to the transdisciplinary field of statistical thermal physics. However, in order for the concept of Shannon's theoretical-communicative entropy to have relevance within this field of physics along the lines suggested by von Neumann (i.e. information from an observer about an observed phenomenon), it would have to appeal indispensably to a notion of an evident semantic character, which would contradict the Shannonian dogma of "semantic irrelevance." It was Weaver himself who opened the door to the possibility of developing a theoretical account of strictly semantic information based on Shannon's proposal "But this does not mean that the engineering aspects are necessarily irrelevant to the semantic aspects."[22]

3. Naturalizing Information in Scientific Practices: Beyond the Non-Semantic Dogma

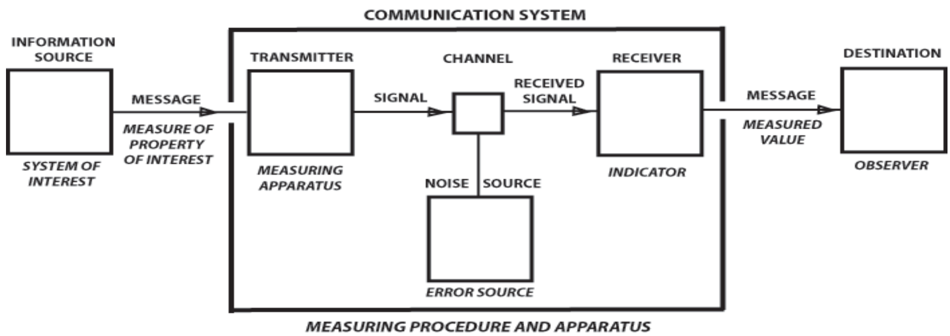
Just after Shannon's paper on communication theory was reissued in the famous book, it became extremely popular within the scientific community of that time. This popularity was accompanied by an enormous interest in the application of his technical proposal in disciplines highly disconnected from the transmission of signals, mainly molecular biology and thermal statistical physics. In the latter case, the choice of the name 'entropy' and the suggestive comments of von Neumann and Weaver played an indispensable role in the progressive informationalization of thermal physics during the 50s. This growing trend encouraged the belief that Shannon's theory would necessarily play a central role in understanding the knowledge-formation process underlying certain scientific practices.

One of the most illustrative examples of this intellectual movement was the paper "Information, Measurement, and Quantum Mechanics" by Jerome Rothstein,

[21] See Shannon and Weaver, *The Mathematical Theory*, note 1.
[22] See Shannon and Weaver, *The Mathematical Theory*, 8.

published only two years after Shannon and Weaver's book.[23] In this paper, the author argued that the process of scientific measurement (for instance, measuring the temperature of a substance) could be understood by Shannon's theory as a process of communication between the observed system as information 'source' (i.e. the temperature of a substance) and the observing agent as 'destination.' In this analogy, (i) the 'message' would correspond to the value of the measured property (e.g. 147º Kelvin degrees), (ii) the 'transmitter' would correspond to the measuring device (e.g. a thermometer), and (ii) the 'receiver' would correspond to the indicator of the measured value (e.g. a scale of values in the thermometer), so that the measuring procedure would act as a communication channel between the observed system and the observing agent. Interestingly, such analogy is conceptually based on assuming the identification (previously suggested by Weaver and von Neumann) between information Shannon H and statistical-mechanical entropy S. Rothstein's analogy[24] showed how to understand this identification: the greater the capacity of the observer to distinguish different microstructures of a system in a measure, the lower the entropy and therefore the lower the information content of that measure. In short, Rothstein argued that Shannon's communication theory would find a direct application in the field of mediated interaction between observers and observed systems, shedding light on the functioning of scientific measurements.

Figure 1. Rothstein's Analogy between Communication and Measuring[25]



[23] See Rothstein, "Information, Measurement," 171-175.

[24] Rothstein's proposal would later influence more sophisticated theoretical informational proposals to scientific measurement such as that of L. Finkelstein ("Representation by symbol systems as an extension of the concept of measurement," *Kybernetes* 4, 4 (1975): 215–223), who recognized that Shannon's formalism only provided the syntactic architecture of the proposal, while measurement theory would add the strictly semantic component. However, many of the deficiencies of his proposal are already found in Rothstein's original paper (Rothstein, "Information, Measurement").

[25] Rothstein, "Information, Measurement," 171-175.

As far as the objective of this paper is concerned, this Rothsteinian naturalist proposal implicitly attempted against the non-semantic dogma that Shannon defended with respect to his proposal. If we analyze in detail this communicative-theoretical model of scientific measurement, the amount of information Shannon transmitted by the values of the measured properties ends up possessing certain semantic properties alien to the original formal apparatus. Firstly, this semantic character implicitly added to the concept of Shannon is reflected in the fact that observer A acquires communicatively or medially information about (i.e., with meaning or referring to) the microstructure of the observed system B. Secondly, scientific measurement practices as communication events end up somehow locating Shannon information in certain mental states of the observer at the end of the day. That is, by reducing all the macroscopically indistinguishable microstructures of a system through particular measurement,[26] not only does the amount of Shannon information compatible with the measurement result decrease, but also the information (in its ordinary, semantic-epistemic sense) that the agent possesses about the microstructure of the observed system. But as Mari points out, this analogy between communication and measurement becomes inconsistent because every message can be known independently of its transmission, while the state of the measured system could not be known independently of its measurement.[27] Therefore, in seeking to apply the Shannon concept of information to scientific practice, Rothstein would imply an implicit and conceptually inconsistent conflation of the theoretical-communicative (non-semantic) technical sense with the usual content of the term 'information' (semantic-epistemic).

Another of the authors who followed and even brought to its final theoretical consequences this intellectual path traced by Rothstein was the French-American physicist Léon Brillouin, who developed a systematic theoretical-informational reformulation of statistical mechanics based on the identification between the

---

[26] As J. Wicken ("Entropy and information: Suggestions for common language," *Philosophy of Science* 54, 2 (1987): 176-193) points out, this way of applying the concept of Shannon information in the field of statistical gas mechanics erroneously assumes that the concept of 'state' in statistical mechanics (distinguishable between 'micro-state,' a set of microscopic-molecular values that determine a system, and 'macro-state,' a set of macroscopic and measurable values on a system) could also be applied in the context of communication theory. This is not possible, precisely because while physical states are sensitive to temperature changes in the environment, theoretical-communicative states are not.

[27] L. Mari, "Notes towards a qualitative analysis of information in measurement results," *Measurement* 25, 3 (1999): 183–192.

technical notion of information à la Shannon and the negative quantities of entropy[28] or 'negentropy:'

> A more precise statement is that entropy measures the lack of information about the actual structure of the system. (...) Since any of these [indistinguishable] different microstructures can actually be realized at any given time, the lack of information corresponds to actual disorder in the hidden degrees of freedom (...) The connection between entropy and information was rediscovered by Shannon, but he defined entropy with a sign just opposite to that of the standard thermodynamic definition. Hence what Shannon calls entropy of information actually represents negentropy.[29]

To defend that the informational measure of Shannon entropy would have a real use in the field of statistical mechanics, this measure should necessarily appeal to the ability of the observer to distinguish observationally between different microscopic structures of physical systems. In this way, Shannon entropy would not simply measure the information encoded in the micro-states of the observed system but would properly measure the lack of semantic-epistemic information that an observer possesses "about the actual structure of the system." At this point we defend, following Earman and Norton,[30] that Brillouin implicitly employed the technical notion of Shannon entropy in the physical context of statistical mechanics as the usual (and therefore semantic) meaning of the notion 'information.'

As we have just seen in Rothstein's and Brillouin's cases, this tendency to implicitly semantize (note that none of them explicitly defends an alternative semantic-sensitive concept) Shannon's non-semantic informational measure in order to apply it in scientific contexts was particularly remarkable during the first years of popularization of communication theory. One of the first authors to become aware of this phenomenon was the philosophers Rudolph Carnap and Yehoshua Bar-Hillel, who were developing during that time a strictly semantic information theory based on their program of inductive logic. In the presentation of their theory at the famous Macy Conferences (series of lectures organized since the early 1950s about areas surrounding cybernetics and information theory) around 1951, these authors pointed out how many authors of the time exploited the fashionable Shannon entropy to their advantage as if it were a function sensitive to meaning or semantic content:

---

[28] Brillouin, *Science and Information*.

[29] Brillouin, *Science and Information*, 160-161.

[30] John Earman and John Norton, "Exorcist XIV: the wrath of Maxwell's Demon. Part II. From Szilard to Landauer and beyond," *Studies in History and Philosophy of Modern Physics* 30, 1 (1999): 1–40.

> It has, however, often been noticed that this [semantic] asceticism is not always adhered to in practice and that sometimes semantically important conclusions are drawn from officially semantics-free assumptions. In addition, it seems that at least some of the proponents of communication theory have tried to establish (or to reestablish) the semantic connections which have been deliberately dissevered by others.[31]

Interestingly, the very name 'information theory' was one of the main sources of confusion when it came to attributing (unintentionally and implicitly) the strong semantic character of the everyday notion of 'information' to Shannon's technical concept. As Kline shows,[32] the adoption of the name 'information theory' to refer to Shannon's communication theory occurred progressively during the first years of popularization due to the previous existence of information theories in the British sphere, such as Fisher's or Gabor's. The confusion generated by that name could be illustrated in Shannon's following comment to one of the attendees of the Eighth Cybernetics Conference in 1951, the anthropologist Margared Mead, who criticized the use of a technical notion of information that was far from its ordinary meaning: "I wanted to call the whole of what they called information theory signal theory," he said later, "because information was not yet there. There were 'beep beeps' but that was all, no information."[33]

It was precisely Shannon himself who most actively defended during the explosion of popularity of his proposal that (1) his theory was fundamentally a formal-syntactic theory about the statistical analysis of signal transmission, so that its extension to other scientific domains would not be at all immediate; and (2) that his particular concept of information (measure of entropy) was a mathematical function whose values were independent of the possible semantic values of the elements on which it is defined. This will not prevent the scientific community during the 1950s from continuing to misapply Shannon's fashionable theoretical apparatus and misinterpret his concept of information to address disciplinary problems that were not at all related to the transmission of signals through communication channels, as we have pointed out in the case of Rothstein and Brillouin. This led to what Shannon himself called the 'scientific bandwagon' after his theoretical proposal, which was set in motion not because of the results obtained in its multiple applications[34] but in order to take advantage of its growing popularity:

---

[31] Carnap and Bar-Hillel, "An Outline."

[32] Kline, *The Cybernetics Moment*, Chapter 3.

[33] This quotation of Shannon can be found in J. Gleick, *The Information: A History, a Theory, a Flood*. (New York: Pantheon Books, 2011).

[34] "The information theory approach to SM has not, to my knowledge, led to any concrete results" (Amnon Katz, *Principles of Statistical Mechanics: The Information Theory Approach* (San

> Information theory has, in the last few years, become something of a scientific bandwagon (…) Our fellow scientist in many different fields, attracted by the fanfare and by the new avenues opened to scientific analysis are using these ideas in their own problems (…) It will be all too easy for our somewhat artificial prosperity to collapse overnight when it is realized that the use of a few exciting words like *information*, *entropy*, *redundancy*, do not solve all our problems.[35]

4. Toward a Technical Concept of Semantic Information to apply in Science

We have just seen how there was an important trend within the information bandwagon in the 1950s to implicitly semantize the Shannon information measure. On the other hand, certain authors (aware of the impossibility of using this statistical concept as a measure of meaning) had the intellectual pretension of developing a theory of semantic information, where a measure was postulated that was sensitive to the meaning of the structures considered. Before evaluating the greater or lesser success of its application in specific scientific fields (paradigmatically, the measurement-observation[36] in statistical mechanics), let us first proceed to explore the main proposals of semantic information.

4.1. Bar-Hillel & Carnap: Statistical Semantic Information

As mentioned in the previous section, Carnap and his collaborator Bar-Hillel developed during the 1950s the first quantitative-statistical theory of semantic information.[37] As mentioned by the authors, this theory was framed within the broad Carnapian project to lay the conceptual foundations of inductive logic as a central methodology in the development of the empirical sciences. Unlike the theoretical proposal of Shannon used by Rothstein and Brillouin to describe the measurement processes in statistical mechanics, the theoretical proposal of Carnap and Bar-Hillel was in fact designed to illuminate how our knowledge works in scientific practices such as measurement. Broadly speaking, one of the main objectives of Carnap and Bar-Hillel's theory of semantic information is to statistically measure how the amount of 'semantic information' (understood as information about a physical property) encoded in the output of an experimental measurement E could alter our initial knowledge by decreasing the space of possible hypotheses H compatible with that scenario. Note that this space of possible hypothesis H would

Francisco: Freeman, 1967), i).

[35] Claude Shannon, *The Bandwagon* (1956), in *Collected Papers*.

[36] In this paper we assume that all scientific measurement is an observational process, but not the other way around.

[37] See Carnap and Bar-Hillel, "An Outline."

be partitioned by a conceptual framework L into different particular hypotheses. For example, the L_{SM} conceptual framework of statistical mechanics partitions the space of possible (non-measurable-observable) hypothetical H microscopic configurations of a substance with respect to the (measurable-observable) temperature values of that substance, so that obtaining 189º Kelvin in a thermometric measurement would reduce the space of possible microscopic configurations to a particular subset of this space. From these theoretical elements it could be formulated its measure of semantic information (represented in the function 'inf') contained in a proposition *i*:

$$\inf(i) = -\log m(i) \tag{2}$$

With this quantitative-statistical measure of the amount of semantic information inf(*i*) proposed by Carnap-Bar-Hillel (defined as the negative logarithm of the number of events m referred by the content of proposition *i*) it would be possible to evaluate the particular conceptual framework L that would maximize the information provided by an experimental data in the selection of the most appropriate hypothesis. For example, from the conceptual framework of L_{SM} statistical mechanics, the same proposition *i* about the temperature of a substance would provide more information than from the framework of L_{TD} thermodynamics: namely $\inf_{SM}(i) > \inf_{TD}(i)$. This is precisely because the number of events (hypothetical microconfigurations of the substance) to which the 189º Kelvin refers in a thermometric measurement are reduced more from the framework of statistical mechanics than from the framework of thermodynamics.

In this line, the authors[38] clearly differentiated between the amount of information 'inf(i)' of a logically structured statement or proposition i, which depended on the negative logarithm of the possible alternatives (as with Shannon entropy); and the content measure 'cont(i)' of a proposition i, which refers to the number of compatible events indicated by i. Thus, while in the domain of the function cont(i) (where i = 'the measured temperature of this substance is 189º degrees Kelvin') the set of microscopic events referred to by the value of 189º K included in proposition i would be included, the domain of the function 'inf(i)' would properly cover the number or quantity of events referred to by this same proposition. Contrary to Shannon's theoretical proposal, Carnap and Bar-Hillel argued that their theory was constitutively based on the evaluation of the semantic content of the vehicles of information as symbols-messages, or properly in their case, declarative sentences or even logical propositions:

> Prevailing theory of communication (or transmission of information) deliberately

[38] See Carnap and Bar-Hillel, “An Outline.”

> neglects the semantic aspects of communication, i.e., the meaning of the messages. This theoretical restraint is, however, not always adhered to in practice, and this results in many misapplications. The theory outlined here is fully and openly of a semantic character.[39]

4.2. MacKay: Descriptive Information

When Shannon presented his communication theory at the first London Symposium on Information Theory in 1950, he realized that a significant number of UK scientists (what some historians such as Kline[40] called the English School of Information Theory) used the term 'information theory' to refer to a disciplinary field that transcended that of Shannon's theory: "They adopted the name Information Theory to refer to a broader concept of information than that held by Shannon."[41] If we have to point out the differential feature of the English theories of information as opposed to Shannon's North American alternative centered on the transmission of symbols, this would undoubtedly be the creation of a concept of information focused on the clarification of scientific practices: "the concept of Information has wider technical applications that in the field of communication engineering. Science in general is a system of collecting and connecting information about nature."[42] In this direction, the physicist Donald MacKay sought to develop (during a period spanning from the early 1950s to 1969, the date of publication of his seminal work) a general theory of information that would account for the system of collecting and connecting information that was scientific practice.[43]

From his early interest in the problem of the objective limits of scientific measurement "the art of physical measurement seemed to be ultimately a matter of compromise, of choosing between reciprocally unrelated uncertainty,"[44] the young scientist MacKay will intellectually evolve towards the question of how we should quantify (he will propose the term 'information quantum,' analogous to the 'atomic propositions' of Wittgenstein's Tractatus[45]) and understand information as it naturally appears in the context of measurement in actual scientific practices. In this way, MacKay's information theory will not be detached from particular physical

[39] See Carnap and Bar-Hillel, "An Outline," 1.
[40] See Kline, *The Cybernetics Moment*, 104-111.
[41] Kline, *The Cybernetics Moment*, 105.
[42] Kline, *The Cybernetics Moment*, 206.
[43] See MacKay, *Information, Mechanism.*
[44] See MacKay, *Information, Mechanism*, 1.
[45] "Although (…) Wittgenstein himself had now repudiated the atomistic approach of the Tractatus, the field of scientific measurement seemed well suited to logical treatment on these" (*Ibid*, 2)

theories, as in the case of Shannon's proposal, but it will otherwise emerge naturally (or be already incorporated) from the physical theory itself employed by the agent to obtain knowledge via measurement processes.

Interestingly, MacKay incorporated into his ambitious information theory (i) the Fisher measure of statistical information or 'metric information' (renaming it 'metron') defined as the reciprocity of a parameter O about a variable X; and (ii) the Gabor measure of physical information based on minimum units of phase-frequency-time volume or 'structural information,' renaming the latter as 'logon.' Both measures of information make up what this author calls 'descriptive information,' due to their role in being used by scientific observers to describe the phenomena observed through certain measurement processes. For example, the possible values of temperature-parameter O that we can measure on a substance provide the agent-observer with certain metric and semantic information about (i.e., descriptively or intentionally) the hypothetical variables of position and molecular velocity that make up the microstructure of a substance. In the same way, all the vectors of position and velocity that determine the exact 'micro-state' of a gas at a particular moment provide the agent-observer with certain structural information that accurately describes the microscopic state of the system.

In addition to the metric (Fisher's) and structural (Gabor's) information, MacKay[46] includes in his ambitious theory the Shannon entropy, redefining it as "selective information," precisely because its value depends on how the units-symbols with which the messages to be transmitted are "selected." Recognizing this informational measure as merely syntactic, this author will also argue in his *Information, Mechanism, and Meaning* that the concepts of metric information and structural information (both grouped under the name of descriptive information) possess a semantic character derived from the intentionality of these agent-observers: "It appears from our investigations that the theory of information has a natural, precise, and objectively definable place for the concept of meaning."[47] However, MacKay seems to suggest (although it is not very clear) that the semantic character of these concepts and measures of information would depend mainly on the capacity of the agent-observers to employ them as representational tools of the objective phenomena.

[46] See MacKay, *Information, Mechanism*.

[47] See MacKay, *Information, Mechanism*, 93.

4.3. Dretske: Semantically-Enhanced Shannon Information

So far, we have explored (i) a concept of semantic information developed in parallel with Shannon's measurement (i.e., Carnap-Bar-Hillel information) and (ii) another concept developed as a semantic alternative to this same information measure during the 1950s and 1960s, i.e., MacKay's descriptive metric and structural information. I will now detail a third strategy to obtain a technically defined notion of semantic information, based mainly on the theoretical-formal supplementation of Shannon's measure in such a way that it is effectively capable of capturing certain types of semantic relations existing in the messages transmitted communicatively. Among all the concepts that we can find within this group, one of the most important and influential within literature is that developed by the philosopher Fred Dretske in the early eighties to develop a naturalistic theory of human perception based closely on Shannon's theory of communication.[48]

Although this author recognizes from the beginning the Shannonian dictum that its entropy measure does not capture the plausible meaning of a sequence of symbols, it is also part of Weaver's suggestion that certain engineering considerations might be relevant to measuring certain semantic relationships (see Section 2). To this end, communication theory should be technically and theoretically supplemented to derive semantic relations from the statistical correlations between the occurrence of symbols and the occurrence of events: "the underlying structure of this [communication] theory, when suitably supplemented, can be adapted to formulate a genuinely semantic theory of information."[49] As we pointed out at the beginning, one of the main reasons why Shannon's communication theory is insensitive to the semantic content of messages is precisely because its measurement entropies not a measure of the amount of information of a particular sequence of signs (as one would expect from a measure of semantic information) but a statistical assembly of possible messages. Dretske proposes to develop a measure I of the information transmitted by an individual event y_i (although the author prefers to speak of 'signals' and 'states of affairs' to refer to the events they refer to and to the events referred to, respectively) to solve this theoretical problem, representing the amount of information that a sequence of particular events y_j transmits about another event or particular state of affairs x_i, thus assimilating the intentional character or 'aboutness' that must characterize any robust semantic information concept. This measure could be formulated by the following mathematical expression:

[48] See Dretske, *Knowledge and the Flow.*
[49] See Dretske, *Knowledge and the Flow*, x.

$$I_{x_i}(y_j) = -\log p(x_i) - H\, p(x_i|y_j) \quad (3)$$

where $x_i \in \{x_i\}$, $i = 1, \ldots, m$; $y_j \in \{y_j\}$, $j = 1 \ldots, n$. This measure would represent[50] the amount of information generated by the existence of certain signal-events x_i about the occurrence of other events y_j, which is determined by the logarithm of the improbability of occurrence of the signal-event x_i minus the H-encoded unexpectedness of the conditional probability distribution between the event referring to x_i and the referred event y_i. Therefore, its measure of semantic information (3) would definitely depend on the non-semantic information measure of Shannon entropy (1), deriving semantic relations from the statistical correlations between the two events. With this measure of the amount of information contained in a single event, Dretske would have the necessary resources to define the semantic information that would be contained in a particular signal S about or about a certain event q:

$$\text{A signal S contains the information that q def} = p\,(q|S) = 1 \quad (4)$$

That is, an event that acts as an S signal (that is, the referenced event) would contain information about another q event (or referenced event) if there is a complete correlation between the S signal and the referenced event. This would capture the modal-counterfactual amount of semantic information; namely, that each time the S signal occurs, it necessarily depends on the fact that the p event to which it refers has also occurred. Dretske explains this modal feature by underlying that the event referred q acts properly as the cause of the signal S, wherein S in turn constitutes the effect of that causal connection. Thus, the fact that a signal S has semantic information about q would be based on the existence of an asymmetric causal relationship between q and S, which for reasons of extension we will not go into further. The key here is that, according to the Dretskean proposal, the semantic information would constitute a statistical property derived from the statistical correlation between events, where its quantity is statistically determined by the conditional probability of (4): while the value 1 would represent that the signal S has the maximum semantic information of which q, the value 0 would represent that the signal S has no information about q.

[50] See Dretske, *Knowledge and the Flow*, Section 2.3 and Christopher Timpson, *Quantum Information Theory and the Foundations of Quantum Mechanics* (Oxford: Oxford University Press, 2013).

5. Against Semantifying Quantitative and Shannon Information

We have just described the main proposals for the development of semantic information concepts in relation to Shannon's non-semantic notion. First, Bar-Hillel and Carnap postulated a measure of semantic information formally identical to Shannon entropy, defined as the logarithm of the improbability of a proposition i. Second, MacKay proposed to reinterpret the representational tools we find in scientific practices as types of 'descriptive information' about the referred phenomenon. Finally, Dretske developed a measure of semantic information based on the statistical correlations between S-signals and state of affairs q that we can extract from Shannon's formalism. Next, our task will be to assess whether these proposals can help to illuminate the knowledge of agents in scientific practices in a more satisfactory way (in terms of conceptual consistency and interpretative coherence) than with Shannon's proposal, as we saw in the case of Rothstein and Brillouin.

5.1. Meaninglessness of Carnap-Bar-Hillel Semantic Information in Real Scientific Practices

The semantic information theory of Carnap and Bar-Hillel[51] was directly conceived to analyze the logical architecture of scientific practices, so in principle we could assume that their concept of information can satisfy this role. Now I will argue that this is not the case, since this proposal presents an important conceptual problem that makes it unsuitable for this task. Namely, as in the case of Shannon's proposal, his measurement of the information content of a proposition i refers to those events that are compatible with that content. That is, a proposition will have the highest degree of semantic-informational content when any possible event will be compatible with (or satisfy) said proposition i, thus we will find what is known as a tautological or trivially true proposition. In the same way, a proposition i will be minimally informational when no event is compatible with its content, thus finding ourselves with a contradictory proposition. However, the fact that the maximum value of the semantic information function cont(i) is found in a tautological proposition could indicate that this proposition is the one that provides more information (in the sense of 'knowledge') about the domain of events considered. But what happens is precisely the opposite. While according to Carnap and Bar-Hillel's proposal, tautological propositions are those with the maximum informational content, they are also those that inform us the least (precisely because of their triviality) about the domain in question. And in the same way with

[51] See Carnap and Bar-Hillel, “An Outline.”

contradictory propositions: although they have the minimum informational content, they are the ones that inform us the most (perhaps in excess) about the domain of events. This is precisely what Floridi calls the 'Carnap-Bar-Hillel paradox' of semantic information,[52] which could be synthesized in the idea that the proposition that has the most information about a domain is the one that informs us the least about this same domain.

As might be expected, this has terrible consequences when it comes to analyzing scientific practices. Let us suppose that we consider the proposition i_1 = 'the molecules of this gas possess a certain position and velocity.' According to the proposal of Carnap and Bar-Hillel, the proposition i_1 would have the maximum informational semantic content with respect to the domain of microscopic configurations of the gas, since virtually any possible microstructure of the gas would be (again, trivially) compatible or satisfy such a proposition. However, this same proposition i_1 would not provide any significant information to the agent about the actual microstructure of the gas, since for any sufficiently competent agent it would be completely trivial or non-informative the fact that the molecules of a gas have a certain position and speed. What would be significantly informative for the agent would be a proposition (with much less informational content than a trivial proposition such as i_1) in which the position and speed of the molecules would be determined. Note that in the same way, a contradictory proposition with the minimum informational content (according to the proposal of Carnap and Bar-Hillel) such as i_2 = 'the molecules of this gas possess and do not possess a certain position and velocity' would provide as little information to an agent as a trivial proposition i_1. In any case, the measurement of the semantic information content of Carnap and Bar-Hillel does not satisfactorily represent the degree of informativity that this semantic content would provide to an agent with respect to the domain of phenomena considered, even describing the behavior of quantities of information in a substantially paradoxical way. For these reasons, the proposal of Carnap and Bar-Hillel does not constitute either the development of a concept of semantic information robust enough[53] to significantly illuminate the dynamics of knowledge in scientific practices.

[52] See Floridi, *Philosophy of Information*, Chapter 3.

[53] Floridi (*Philosophy of Information*) described the theoretical proposal of Carnap and Bar-Hillel ("An Outline") as a weak semantic information theory, precisely because of its inability to be sensitive to epistemically relevant values within a particular context.

5.2. Meaninglessness of Dretskean Semantic Information in Real Scientific Practices

Undoubtedly, Dretske's proposal constitutes one of the most important attempts to obtain a technically sophisticated notion of semantic information. However, I will now show how this would still not be robust enough to account for the dynamics of knowledge in scientific practices. We can conceive of two types of objections to the Dretskean program, namely, technical objections and theoretical objections. As for the former, we can follow Timpson[54] in criticizing the concept of semantic information $Ix_i(y_j)$ (3) as a highly deficient measure of the information that x_i contains about y_i. The central reason is that the first term of the measure $-\log p(x_i)$ (formally identical to Carnap and Bar-Hillel's semantic information measure) would be completely independent of the second term $H\ p(x_i\,|y_j)$:

> For example, our uncertainty in x_i given y_j might be very large, implying that we would learn little from y_j about the value x_i, yet still the amount said to be carried by y_j about x_i, under Dretske's definition, could be arbitrarily large, if the surprise information$_t$ of x_i dominates. Or again, the channel might be so noisy that we can learn nothing at all about x_i from y_j -the two are uncorrelated, no information can be transmitted- yet still $Ix_i(y_j)$ could be strictly positive and very large (if the probability of xi is sufficiently small). This is sufficient to show that $Ix_i(y_j)$ is unacceptable as a measure. The hoped-for link to information theory is snapped.[55]

That is, let us assume in a context of scientific practice that events y_i represent the possible microscopic (non-observable) configurations of a gas, as well as events y_i represent the possible observable temperature values (e.g., marks on a thermometer) that we can extract by thermometric measurement from a gas. In this scientific measurement scenario, an agent could possess a high Shannon-encoded uncertainty $H\ p(x_i\,|y_j)$ on the actual microstructure of the gas y_i given a set of microconfigurations compatible with a particular temperature value x_i, which represents that the agent cannot recognize which is the actual microstructure only from her knowledge of its temperature value (i.e. 189º Kelvin). At the same time, according to Dretske,[56] the amount of information that the 189º Kelvin of the gas contains about the actual microstructure of the gas can be arbitrarily high, and therefore it would be independent of the uncertainty that the agent possesses about such microstructure.[57] Consequently, the amount of semantic information $Ix_i(y_j)$

[54] Timpson, *Quantum Information Theory*, 40.

[55] Timpson, *Quantum Information Theory*, 40.

[56] See Dretske, *Knowledge and the Flow*, x.

[57] "Unfortunately, the quantity $Ix_i(y_j)$ cannot play the role of a measure of the amount of information that y_j carries about x_i. To see this we need merely note that the surprise [Shannon] information$_t$ associated with x_i is largely independent of the uncertainty in the conditional

could not satisfactorily capture the amount of information that the thermometric values of the gas carry on the actual microstructure of that same gas, and with even less technical consistency would represent the amount of informativeness of the thermometric signal for the competent agent with respect to the molecular properties of this substance.

As for the theoretical deficiencies that we could object to Dretske's proposal of 'information that', we will find its excessive idealization of the informational dynamics. In the first place, in order to specify how much information content would be linked to the occurrence of the event we should determine the domain of different possible events. Illustratively, to know how much information the 189º K event provides us about the actual microstructure of the gas we should know in advance how many different possible microstructures are compatible with that thermometric signal. However, it would be reasonable to think that it is not possible to specify the domain of different events in realistic contexts of application of the Dretskean apparatus. In fact, if we consider the measurement practices of classical statistical mechanics, the fact that the number of distinct micro-stages of a gas is uncountably infinite (precisely because the classical phase space of a gas is continuous) makes it practically and conceptually impossible to determine the domain of events in this particular case. Therefore, Dretske's measure will not only be technically deficient because of its combination of two independent terms in (3), but also theoretically ideal, since it could only be applied in extremely simple and uninteresting scenarios with respect to our understanding of sufficiently realistic scientific practices.

6. Defending a Neo-MacKayian Meaningful Use of Information Concepts in Science

Unlike the proposals of Carnap-Bar-Hillel and Dretske that we have previously analyzed, MacKay's theory of semantic-descriptive information does not depend on the development of a measure of its own, but on a kind of interpretative exercise carried out on existing representational structures (i.e. reciprocal quantities of variance, volume of phase space) within real scientific practices.[58] This fact may be a vice or a virtue: a vice if the interpretative exercise is not sufficiently robust or well-defined theoretically (as is unfortunately the case with MacKay), and a notable virtue because of the enormous descriptive potential that a well-defined proposal developed in this direction would possess. In this section I will attempt to carry out this task.

probability distribution for x_i given y_j" (Timpson, *Quantum Information Theory*).

[58] MacKay, *Information, Mechanism*.

The ambitious information theory that MacKay originally presented at the First London Symposium on Information Theory was severely criticized by his contemporaries during the 1950s, right in the middle of the maelstrom of Shannon's bandwagon. As Kline reminds us, Colin Cherry (also from the English School of Information Theory) argued in 1956 that Bar-Hillel and Carnap's semantic information theory was the “only investigation of which your author is aware, into the possibility of actually applying a measure to semantic information,” an implicit critique of MacKay.[59] This disregard for MacKay's information theory, with no repercussions in the literature, could be justified on a theoretical level by the enormous indefiniteness of his proposal and by the vagueness of his definitions, as I will proceed to show. As far as we are concerned in this paper (i.e., the application of informational concepts in mediated practices in science), the author states that “in physics, of course, the notion of descriptive information-content is often more useful than that of selective [Shannon] information-content. A physicist may want to know how much theoretically available information is being wasted in a given microscope.”[60] As this quotation shows, MacKay defends that his concept of descriptive information (Section 4.2) is a truly significant alternative to the informational concept of Shannon entropy (on which the proposals of Carnap-Bar-Hillel and Dretske, respectively, rest indirectly or directly) to be applied in an epistemically useful way in the context of scientific practices.

But what does the author mean by 'theoretically available information' about a physical system? If we analyze the text in depth, we can arrive (seventy-seven pages later) at something similar to a definition: “A particular message may now be pictured as selecting a particular region, which may be identified by the vector (or a distribution of vectors) linking it to the origin. The meaning of the message is then represented by the orientation of this vector, relative to the vector basis.”[61] Therefore, from this proposal the semantic information that the agent-observer has on the observed system is a resource that is defined by the theory used, which determines what MacKay calls the 'representational space' built by the vectors that parameterize the system under consideration. However, here we argue that it is not necessary to introduce an additional space to that already used in real scientific practice. In the case of statistical mechanics that we have used throughout this paper, the phase space of the system itself (abstract space that encodes all the possible values of position and molecular speed of the system) would play the role of

[59] See Kline, *The Cybernetics Moment*, 111.
[60] See MacKay, *Information, Mechanism*, 15.
[61] See MacKay, *Information, Mechanism*, 92.

representational parameterization of statistical-mechanical information about the system.

The main flaw of MacKay's proposal lies precisely in seeking to quantify the meaning of a description, represented in a vector within the phase representational space, by means of the orientation of this vector. That a vector-description has a certain orientation is a conventional-arbitrary fact (subject to how we design the representational space) and in no way exhausts the semantic information encoded in such a description within the system's representational space. If this were the case, then after performing a thermometric measurement we would obtain a new statistical-mechanical description whose meaning would be the opposite of the description prior to the thermometric measurement or observation, precisely because the rules of vector calculus state that a difference in vectors changes the orientation of the resulting vector (see Figure 2).

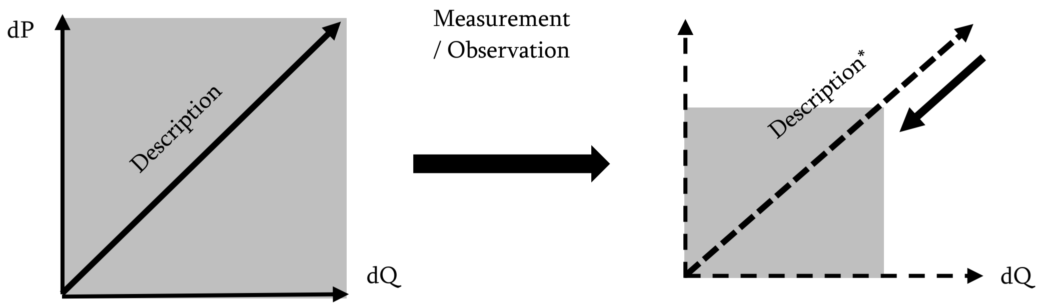


Figure 2. MacKay semantic information defined in representational space. The phase space represents information about the molecular position (P) and velocity (Q), where the orientation of the vector-description represents its meaning. Therefore, according to the rules of vector calculation, the meaning of the description* (or its semantic information) after performing a measurement on the system would be the opposite of meaning of the previous description.

However, it would be absurd to think that the semantics of this sort of descriptions made in a realistic measurement scenario closely depends on the rules of vector calculus: these rules fix how we can construct vector representations of a phenomenon, but they do not properly fix the meaning of those theoretical elements we use to describe them. Contrary to MacKay's concrete proposal but following his framework, we propose to understand the semantic information of a statistical-mechanical description as the way in which theory delimits the way in which representational resources (i.e. micro-states as 'phase points', macro-states as 'phase regions') represent target phenomena. Illustratively, statistical mechanics[62] (not vector calculus or any other mathematical theory employed) theoretically determine (i) that each point of the representational phase space encodes exact semantic information about the position and velocity of each and every molecule that composes a particular substance; and (ii) that each region of the representational phase space would encode semantic information about the macroscopic properties (temperature, pressure, and so on) on this substance.

It is the theoretical apparatus of statistical mechanics that specifies how these two types of semantic information (non-observable information about molecular components, observable information about substances) are interconnected in a context of scientific practice such as measurement; and in no case could this be specified in a realistic formal-theoretical apparatus of the statistical theory of signal transmission (against Rothstein[63] or Dretske[64]). For example, in Boltzmann's statistical-mechanical formalism, each piece of observable semantic information (e.g., the 189º K value) would be theoretically associated with a counterfactual set of non-observable semantic information (e.g., semantic information about two observationally indistinguishable microscopic configurations). In this way, an agent competent with the practices of statistical mechanics could extract significant semantic information about the real micro-state of the system from the macroscopic values of observable properties (e.g. temperature) measured, the latter associated with sets of micro-state counterfactually compatible with that macroscopic value.

Another of the main advantages of our Neo-MacKayian proposal is that it would allow us to incorporate the theory of scientific measurement recently defended by van Fraassen, where the qualitative concept of semantic-epistemic information would play an essential role. For van Fraassen[65], the possible measurable states of a physical system are represented in a parameterized space. Thus, a particular measurement would inform us about the state of the target system by reducing the parameterized space of possible values to a subsystem. In our case of study of the measurements in statistical mechanics, the parameterized space is none other than the phase space of possible microscopic configurations of the system, which would be reduced to a particular sub-region (sub-set of micro-states) after a macroscopic measurement on the system (see Figure 2). In this way, the meaning of

[62] Note that to defend this particular thesis we do not need to support any particular conception of scientific theories.
[63] Rothstein, "Information, Measurement."
[64] See Dretske, *Knowledge and the Flow*, x.
[65] Bas van Fraassen, *Scientific Representation: Paradoxes of Perspective*, (Oxford: Oxford University Press, 2008), 141-185.

the semantic information about a gas contained in the phase space would not change radically (as suggested by MacKay[66]) after a thermometric measurement on this gas, but would increase its epistemic relevance for the agent-observer with respect to the recognition of the microscopic configuration of the measured system.

Unlike the quantitative-semantic information proposals of Carnap-Bar-Hillel and Dretske, MacKay's information-theoretical framework (although not so much his particular proposal) would allow us to develop a robust naturalistic theory about how scientific agents obtain semantic information without the need to postulate forced analogies with signal transmission processes or without relying on an information concept formally close to Shannon's. Our proposal has been an apology for the use of the qualitative concept of 'semantic information' within a robust philosophical analysis of scientific practices (in our case we have focused on measurement) against the hegemonic tendency since the mid-twentieth century to quantify-model the semantic information of scientific agents with the formal basis of Shannon's communication theory.

7. Conclusion

In this paper we have shown how the first attempts to apply the concept of Shannon information in the field of scientific practice (e.g., Rothstein's and Brillouin's[67]) failed precisely because of the lack of semantic character of this notion. Therefore, we have explored the main proposals of semantic information theories developed on Shannon's quantitative notion applied in the context of science (e.g. Carnap and Bar-Hillel's, and Dretske's[68]), arguing later that each of them lacks the theoretical capacity to satisfactorily account for the information dynamics between observers and systems observed in realistic scientific contexts. Finally, we have defended not the particular proposal but the informational framework of MacKay,[69] based on the informational interpretation of the representational resources of real scientific practices and arguing that it would be robust enough to be satisfactorily employed in philosophical analyses of this field. The following quote from Colin Cherry clearly shows the spirit of what we have sought to defend throughout this paper:

> Questions of extracting information from Nature and of using this information our models or representations lie outside communication theory - for an observer looking down a microscope, or reading instruments, is not to be equated with a

[66] MacKay, *Information, Mechanism*, 92.

[67] See Rothstein, "Information, Measurement, and Quantum Mechanics," *Science* 114 (1951): 171-175 and Brillouin, *Science and Information*.

[68] Dretske, *Knowledge and the Flow,* x.

[69] MacKay, *Information, Mechanism.*

listener on a telephone receiving spoken language. Mother Nature does not communicate to us with signs or language. A communication channel should be distinguished from a channel of observation.[70]

[70] E.C. Cherry "On Validity of Applying Communication Theory to Experimental Psychology," *British Journal of Psychology* 48 (1957): 176-188, quotation in Kline, *The Cybernetics Moment*, 111.