Logic, Epistemology, and the Unity of Science 52

Claudia Fernández-Fernández

Awareness in Logic and Epistemology

A Conceptual Schema and Logical Study of The Underlying Main Epistemic Concepts

Springer

Claudia Fernández-Fernández
Departamento de Filosofía
Universidad de Málaga
Málaga, Spain

ISSN 2214-9775 ISSN 2214-9783 (electronic)
Logic, Epistemology, and the Unity of Science
ISBN 978-3-030-69605-4 ISBN 978-3-030-69606-1 (eBook)
https://doi.org/10.1007/978-3-030-69606-1

This Springer imprint is published by the registered company Springer Nature Switzerland AG
The registered company address is: Gewerbestrasse 11, 6330 Cham, Switzerland

Contents

About the Author

Claudia Fernández-Fernández is graduated in Philosophy (2011) by the University of Málaga (Spain), Master in Analytic Philosophy (2013) by the University of Barcelona (Spain) and Ph.D. in Philosophy (2019) by the University of Málaga (Spain), with a dissertation on Knowledge and Awareness in Logic and Epistemology. Her research interests are related to Logic, Philosophy of Logic, Philosophy of Mathematics and Philosophy of Language.

Since 2016 she works as a substitute interim teacher at the Department of Philosophy in Málaga University (Spain), teaching courses related to Philosophy of Logic, Philosophy of Mathematics and Semantics. Most of her publications are related to the topic of her Ph.D. dissertation, such as Fernández-Fernández, C. and Velázquez-Quesada, F. R. (2020). Awareness-of and awareness-that their combination and dynamics. *Logic Journal of the IGPL*, DOI: https://doi.org/10.1093/jigpal/jzz043; Fernández-Fernández, C. (2019). Awareness Logic: an Epistemological Defence. Correlations between Awareness Logic and Epistemology, *Kairos. Journal of Philosophy &Science*, 22(1), 72–85. DOI: https://doi.org/10.2478/kjps-2019-0009; and Burrieza, A. and Fernández-Fernández, C. (2015). Distributed Explicit Knowledge and Collective Awareness. *Bulletin of the European Association for Theoretical Computer Science*, Number 116 (June 2015), pp. 195–203.

She has presented her work on several international conferences such as *XX Model-Based Reasoning in Science and Technology. Inferential Models for Logic, Language, Cognition and Computation*, October 2018 in Seville (Spain); *The 18th Trends in Logic International Conference*, September 2018 in Milan (Italy); *Logica 2018*, June 2018 in Hejnice (Czech Republic); *Logica 2017*, June 2017 in Hejnice (Czech Republic) and *Third Lisbon International Conference on Philosophy of Science: contemporary issues*, December 2016 in Lisbon (Portugal).

Abbreviations

AGM	Alchourrón, Gärdenfors and Makinson
AI	Artificial Intelligence
AJI	Awareness Justification Internalism
AL	Awareness Logic
ANM	Awareness Neighbourhood Model
DEL	Dynamic Epistemic Logic
EAK	Explicit Aware Knowledge
EL	Epistemic Logic
JL	Justification Logic
JTB	Justified True Belief
LP	Logic of Proof
MWD	Merriam-Webster Dictionary
OD	Oxford Dictionary
PAL	Public Announcement Logic

List of Figures

Chapter 1
Introduction to Epistemic Logic and Epistemology

Abstract This chapter summarizes the most relevant preliminary concepts from Epistemic Logic and Epistemology. Regarding the former, the problem of logical omniscience and its different solutions are revised, with Awareness Logic as the chosen solution for carrying out this research. With respect to Epistemology, a brief historical background is presented and the view of Awareness Justification Internalism is highlighted as the one that will serve as a theoretical background for the logical approach. To conclude, some aspects of the disconnection between both disciplines are pointed out and the concept of Epistemic Awareness is defined as the bridge-notion through wich the theoretical re-connection will be developed.

Keywords Epistemic logic · Epistemology · Epistemic internalism · Awareness logic · Logical omniscience

1.1 Epistemic Logic: Main Concepts

This section provides the reader with a brief and sometimes schematic presentation of the main logical concepts that will be employed during the following chapters. While introducing these main notions some of them will be highlighted as significantly relevant; and in those cases where it comes down to different interpretations some choices will be made and justified.

Epistemic Logic. Epistemic Logic (EL) was 'born' in von Wright (1951). Von Wright took the first steps towards a formal study of knowledge and belief. Some years later, in Hintikka (1962), the author transforms von Wright's ideas into more semantic concepts and EL was established as what we know today. EL has been developed and revisited by many authors since then.

Epistemic Logic[1] is an extension of Modal Logic, in which the necessity operator ($\Box$) is interpreted as knowledge and ranges over a set of epistemic agents ($\mathtt{Ag}$),[2]

[1] What follows is one possible contemporary presentation of EL.

[2] In *doxastic logic*, the logic of belief, the necessity operator is interpreted as what the agent believes.

C. Fernández-Fernández, *Awareness in Logic and Epistemology*,
Logic, Epistemology, and the Unity of Science 52,
https://doi.org/10.1007/978-3-030-69606-1_1

such that the formula $\Box_i\varphi$ is read as *agent* i *knows* φ (or 'agent i knows the truth of φ' or 'agent i knows that φ is the case'). The semantic structure of EL, like Modal Logic, is based on the *possible worlds model* (also called Kripke model), where every epistemic possibility (possible world) an agent may consider is represented via the epistemic accessibility relation of each agent to those possible worlds she considers.

An **epistemic model** is the structure in which the knowledge of the agents is represented. Let $\mathtt{P}$ be a set of atomic propositions and let $\mathtt{Ag}$ be a set of epistemic agents. Then an epistemic model is a tuple $\langle W, \{R_i\}_{i\in\mathtt{Ag}}, V\rangle$, such that:

- $W \neq \varnothing$ is a non-empty set of possible worlds,
- $R_i \subseteq (W \times W)$ is the epistemic accessibility relation of agent i (indicating the worlds each agent considers possible from each one of them), and
- $V : W \to \wp(\mathtt{P})$ is the atomic valuation function (indicating the atomic propositions in $\mathtt{P}$ that are true at each possible world).

The pair (M, w), with $w \in W$, is called an **epistemic state in model M** and corresponds to the evaluation point, that is, the world in which the given formula for the given agent will be assessed based on the other worlds that are accessible from w for agent i. Concretely, from each world $w \in W$, the agent considers possible all those worlds u that she can R-access from w.

The epistemic model M is described by the **language** of epistemic logic, $\mathcal{L}(\Box)$, whose formulas φ, ψ are given by:

$$\varphi ::= p \mid \neg\varphi \mid \varphi \wedge \psi \mid \Box_i\varphi, \quad \text{where} \quad p \in \mathtt{P} \,\text{and}\, i \in \mathtt{Ag}.$$

Recall that $\Box_i\varphi$ is read as *agent* i *knows that* φ *is the case*. The other Boolean connectives can be defined from negation ($\neg$) and conjunction ($\wedge$) in the standard way.

The **semantic interpretation** of the formulas is such that for (M, w) being an epistemic state in $M = \langle W, \{R_i\}_{i\in\mathtt{Ag}}, V\rangle$ and φ and ψ any formulas in $\mathcal{L}(\Box)$; $(M, w) \Vdash \varphi$ indicates that φ is true at w in M (and $(M, w) \nVdash \varphi$ indicates that it is not true). Then,

- $(M, w) \Vdash p$ iff $w \in V(p)$
- $(M, w) \Vdash \neg\varphi$ iff $(M, w) \nVdash \varphi$
- $(M, w) \Vdash \varphi \wedge \psi$ iff $(M, w) \Vdash \varphi$ and $(M, w) \Vdash \psi$
- $(M, w) \Vdash \Box_i\varphi$ iff for every $u \in W$, $R_i wu$ implies $(M, u) \Vdash \varphi$.

Note how the semantic interpretation of $\Box_i\varphi$ says that agent i knows φ at w if and only if φ is true at all worlds u that are accessible for i from w due to R_i.

The concepts of **satisfiability** and **validity** are defined in a standard way. Let φ be a formula in $\mathcal{L}(\Box)$. A formula is said to be *satisfiable* if and only if it is true in at least one world w of at least one model M (that is, if and only if there is one M such that $(M, w) \Vdash \varphi$). A formula is said to be *valid* if and only if it is true at all worlds w of all models M. The valid formulas are represented as follows: $\vdash \varphi$.

The **knowledge operator** ($\Box_i$) stems from the universal necessity modal operator. As such, it has the same properties as the modal operator. This means that the knowledge operator satisfies both the *rule of necessity (Nec)*, $\vdash \varphi$ implies $\vdash \Box_i\varphi$, and the

K-axiom, $\vdash \Box_i(\varphi \to \psi) \to (\Box_i\varphi \to \Box_i\psi)$. The first one states that if a formula is true, then agent i knows it. The second property makes the knowledge closed under logical consequence, meaning that if agent i knows φ and $\varphi \to \psi$ at world w, then she automatically knows ψ via the *Modus Ponens* rule. And this, conversely, means that now ψ is also true at every u accessible from w. This feature of Epistemic Logic has been given the name of the *problem of logical omniscience*.

The problem of logical omniscience. The fact that the knowledge of the epistemic agents that EL proposes is closed under logical consequence should not be a problem by itself. The controversy arises when the notion of 'epistemic agents' is applied to human beings, instead of abstract entities that posses knowledge (like some ideal agents from Artificial Intelligence (AI)). But, of course, when one tries to understand the concept of 'epistemic agent', the most common and intuitive interpretation is to relate it to ourselves.

This was also the case of Hintikka, who in his seminal work argued in favour of Epistemic Logic as a way of representing human knowledge. He was already aware of the difficulties that this would generate and stated that: "[t]he logical implications of what we know do not come to us without any work on our own part; they are truths which we can extract, often with considerable labor, from whatever information we already have" (Hintikka 1962, p. 37). He continued advocating that although the laws of logic are not equivalent to the laws of thought, they could be the "laws of the sharpest possible thought" (Hintikka 1962, p. 37). A few years later, in (Hintikka 1975) he proposed one strategy to solve the logical omniscience of epistemic agents as will be explained below.

Human subjects, though rational, are not omniscient (we could leave this adjective for the knowledge of some God(s)). In fact, humans, as computing machines, have *limited reasoning abilities*. Though the rationality of human subjects has not been questioned, what is clear is that it is limited, as our lives are.

In *Reasoning about Knowledge* Fagin et al. (1995), the authors devote a chapter to the problem of logical omniscience (LO) and its possible solutions.[3] They state that the problem as defined above corresponds to *Full Logical Omniscience*. There are also weaker forms of omniscience derived from the Full LO, such as the closure under logical implication or closure under logical equivalence. Other types of LO, not derivable from the Full LO are also mentioned, like closure under conjunction or closure under valid implication. As may already be clear, all forms of LO are related to some kind of closure property that is imposed on the knowledge of the agents in Epistemic Logic.

As the authors wisely conclude, the different strategies out of LO pass through a reconsideration of the conditions for knowledge or the definition of 'knowledge' itself. In the mentioned chapter (Fagin et al. 1995, pp. 309–362) they make a thorough analysis of the existing solutions to the problem of LO by that time, classifying them with respect to the semantic or syntactic structure that is imposed and the different concepts that are being redefined. A complete overview of these solutions may be

[3]For an updated overview and contemporary proposal about this topic see Hawke et al. (2019).

found in the given reference and would lie beyond the limits of this chapter. Some key notions for a general glimpse should thus suffice.

The **solutions to the problem of LO** need to reconsider the conditions for knowledge to be true. In standard EL knowledge is true at a given world if it is true in all worlds accessible from the given one, and that was exactly the cause of the problem. It is an ideal representation of knowledge. One way of solving the problem would be to establish a *different definition of knowledge*, from a syntactic or a semantic point of view. Though attractive as a solution to the problem, these strategies may also leave knowledge without other properties that were important; they are useful for defining knowledge, but not for modelling it formally.

There are also many *non-standard logics* that redefine knowledge or the whole structure of Epistemic Logic. These non-standard approaches commonly change the notion of truth by extending some operator or some part of the semantic structure with a counterpart where not every property or validity holds. Though interesting and applicable to concrete problems in AI and other fields, the authors do not highlight any of them as a complete solution to the problem of LO (if such 'complete solution' even exists).

The two most promising strategies they present, from an intuitive and theoretical point of view, are *impossible worlds* and *awareness logic*. The impossible worlds approach was proposed in Hintikka (1975). He assumes the problem of LO and proposes to change the semantic structure, inspired by contemporary ideas from Veikko Rantala on surface models in Rantala (1975). Hintikka establishes a difference between an *epistemic alternative* and a *logical possibility*, where the former are all the alternatives the agent may consider (including impossible worlds in which, for example, contradictory information may hold), and the latter are maximally consistent descriptions of the world. Using a complex structure he manages to adapt the semantics of possible worlds for including also impossible worlds, having now the possibility of evaluating formulas in both types of worlds. It can be considered a good solution for the problem of LO, since agents are not logically omniscient with respect to the impossible worlds, but, on the other hand, they can 'know' everything in those worlds, making their knowledge in impossible worlds quite trivial.

I left **Awareness Logic** (AL) until last, since it is the strategy I consider most suitable for the purpose of solving the problem of LO in EL. The main idea behind this approach, already proposed in Fagin and Halpern (1988), is that in order to know something, it is necessary to be aware of it. They state it in the following terms:

> The underlying idea is that it is necessary to be *aware* of a concept before one can have beliefs about it. One cannot know something of which one is unaware. Indeed, how can someone say that he knows or doesn't know about p if p is a concept of which he is completely unaware? (Fagin et al. 1995, p. 337. Original italics).

With this view in mind, they propose to add awareness to the standard system of EL as a syntactic operator, acting as a filter on the knowledge of the agent. They also incorporate a new operator for the *explicit knowledge*, representing the 'real' knowledge the agent possesses. The explicit knowledge of the agent is then defined

by the conjunction of both implicit knowledge (standard knowledge in EL) and awareness.

Awareness is defined as an arbitrary list of formulas and only those that coincide with the ones implicitly known, will be the explicitly known formulas. Formally, they modify the original language $\mathcal{L}(\Box)$, being now $\mathcal{L}^A(\Box)$, formed by $\mathcal{L}(\Box)$ and two new operators: the *awareness operator*, A, such that formulas of the form $\mathrm{A}_i\,\varphi$ are read as *agent i is aware of* φ; and the *explicit knowledge operator*, X, with formulas of the form $X_i\varphi$ read as *agent i knows* φ *explicitly*.

The semantic model M is modified, adding the *awareness function*, $\mathfrak{A}$, that assigns a list of formulas, of which each agent is aware of, to each world. Then M^A is an *awareness model*: $M^A = \langle W, \{R_i\}_{i\in \mathtt{Ag}}, V, \mathfrak{A}_i\rangle$.

The *semantic interpretations* for the two new operators are the following ones: let (M^A, w) be an awareness epistemic state with $w \in W$ and φ any formula in $\mathcal{L}^A(\Box)$, then

- $(M^A, w) \Vdash \mathrm{A}_i\,\varphi$ iff $\varphi \in \mathfrak{A}_i(w)$
- $(M^A, w) \Vdash X_i\varphi$ iff $(M^A, w) \Vdash \mathrm{A}_i\,\varphi$ and $(M^A, w) \Vdash \Box_i\varphi$

From the latter, the authors conclude the following validity, which is the key of this proposal: $\vdash X_i\varphi \leftrightarrow (\mathrm{A}_i\,\varphi \wedge \Box_i\varphi)$, stating the logical equivalence between agent i having explicit knowledge about φ and her fulfilling both implicitly knowing φ *and* being aware of φ.

The original proposal leaves the awareness function without any restrictions, but they argue that different closure properties might easily be achieved, such as closure under subformulas or self-reflection. This proposal does overcome the problem of LO and maintains the original structure of EL. It is also compatible with the previous proposal of impossible worlds and with some other semantic structures. Probably the stroke of genius that comes with this proposal, is that Fagin and Halpern managed to formalize a common sense intuition that was already present in many other fields of study.

The potential that AL has, was fruitfully explored in the following decades and incorporated to the new paradigm of the *studies of language and information*. There are also other proposals that include an external factor to the main structure of EL in order to divide knowledge into implicit and explicit and hence prevent the agents from being logically omniscient, but I will not review them here for lack of interest to this research. Though the authors present it as one possible solution among others, they had already explored this approach (as mentioned before) and continued afterwards this line of research.

New approaches of knowledge and information. The last decades have been very prolific in the studies of language and information from the logical point of view. One could speak of a new *paradigm* that has been settled regarding the logical interactions of language and information (and also computation).

The different ways in which information is represented will give rise to the different logics and semantics that one finds nowadays. The milestone that turned the

concepts from Epistemic Logic into a whole new paradigm was probably the concept of *information change*. In contemporary approaches it is hard to find the mere concepts of 'knowledge' and 'belief' as such. Instead, one finds the notion of 'information' and its various interpretations will turn it into either knowledge, or belief, or other types of epistemic attitudes. Of course the concept of 'information change' has its origins in the studies of communication, that are seen as multi-agent scenarios in Epistemic Logic, where the given formulas or propositions change during the communication process.

Epistemic Logic represents a simple semantic interpretation of information. The new approaches to the notion of information (change) have been made from a syntactic and semantic point of view. The purely *syntactic representations* understand 'information' as a list of formulas, or propositions, and sometimes add some properties to the set of information that is assigned to an agent for achieving (or coming closer to) omniscience. The interactions between these formulas or propositions are then deductive. There are different proposals, with *Belief Revision* being one of the most relevant ones (for a general overview see Gärdenfors (1992) and Williams and Rott (2001), for example).

In **Belief Revision Theory** one could highlight the *AGM-model* (named after the authors Alchourrón, Gärdenfors and Makinson that published the originating paper Alchourrón et al. 1985). This proposal of representation of belief establishes a distinction between the *belief set* (consistent set of formulas closed, normally, under logical consequence) and the *belief base* (simple set of formulas acting as a basis for the belief set, without closure properties). The AGM-model provides tools for formalizing the different types of changes (e.g., revisions and updates) there might be in the agent's belief set, but from a purely syntactic point of view. As pointed out in (Velázquez-Quesada 2011, p. 9), "syntactic approaches have been criticized as being too fine-grained, making differences in meaning where there seems to be none". However, Belief Revision Theory had great influence in the forthcoming approaches and is still very useful in different fields of Artificial Intelligence.

The *semantic approaches* in this new paradigm are also based on the idea that information changes, and so do the assigned truth-values; and this change needs to be reflected by the formal structure. The *dynamics of information* is the most important concept that arises with this considerations. The fact that the new approaches adopt a dynamic perspective implies that information is not a static concept any more, instead it is in constant motion. What brings the informational entities to move are the *epistemic actions*, that are formalized as operators that change the model. This branch of logics classifies the actions that promote the information to change.

Dynamic Epistemic Logic. The designation of *Dynamic Epistemic Logic* (DEL)[4] includes different approaches, sharing as a common ground Epistemic Logic as their foundation and epistemic actions as its modification. One of the most salient handbooks regarding the dynamics of information is van Ditmarsch et al. (2008). In

[4]The first research that considers the appearance of DEL a change of paradigm is Gochet (2002). DEL as it is considered here is best developed in van Benthem (2011).

there, the authors define very precisely in which sense the concept of 'information' is now interpreted.

> We regard information as something that is relative to a subject who has a certain perspective on the world, called an *agent*, and the kind of information we have in mind is meaningful as a whole, not just loose bits and pieces. This makes us call it *knowledge* and, to a lesser extent, *belief*. This conception of information is due to the fields known as *epistemic* and *doxastic logic* (van Ditmarsch et al. 2008, p. 1. Original italics).

As mentioned above, DEL serves as a designation for different approaches and logics, inspired partly by EL and partly by Belief Revision theory. The different logics and their interactions that have been proposed in the last decades conform a wide range of possibilities. I will mention the most important notions that appear in those proposals and that make DEL constitute a new paradigm.

The most relevant notion is that of *epistemic actions*, understood as the fulfilment of an epistemic act that causes the existing information of the agent to change by loosing information, gaining new pieces of information or transforming it. The first and most simple action that was formalised is the act of communicating new information, captured in **Public Announcement Logic** (PAL) (see Plaza 1989, 2007). I called it 'simple' since this logic formalises an idealisation of the process of communication, i.e. it represents informational changes "per the occurrence of completely trustworthy, truthful announcements" (Baltag and Renne 2016). This logic deals with how the announcement of new information alters the existing one, changing part of what the agent had and creating new information derived from the old one.

Different logics that include different types of information loss or wins and updates have been developed. The variety of epistemic actions correspond to a huge classification that keeps on growing with every new research. Of course, as happened with standard EL, one also finds here the problem of logical omniscience and some actions are designed to prevent the agents from different types of omniscient properties, while other maintain some closure properties and have been applied to Computation, Artificial Intelligence or Economics.

The main feature that defines non-omniscient agents is to have **limited reasoning abilities**. It is not intended that agents cannot reason at all, but to prevent full logical omniscience the agents each logic presents need to have some kind of limiting mechanism that makes their inferential processes not infinite and provides them with just enough rationality for creating new information by themselves, but at the same time prevents them from being able to be informed about everything there is.

In line with this, there have been some very interesting proposals that, based on Awareness Logic and adopting a dynamic perspective, incorporate the actions of *becoming aware* (and also unaware) or *performing a deductive inference*. Both are individual actions (only affecting one agent) that help limit the deductive possibilities of the agent, while reconciling itself with the common sense. These proposals mix different syntactic and semantic structures in order to be able to capture the intended consequences of each action.

To sum up, one could understand the new paradigm of the dynamics of information as an interdisciplinary field of study, where logicians, linguists, computational

engineers and philosophers have established as a common ground the concept of 'information change', and constructed a new branch of knowledge. Contemporary advances in the field are constant and new interpretations of the basic concepts come with almost every new proposal that is conceived.

1.2 Epistemology: An Overview and a Choice

This section starts by introducing the concept of *Epistemology* and its historical origins. After describing some of its basic ideas it explores the contemporary picture of Epistemology and reviews some classifications. Then, one specific approach, Awareness Justification Internalism, is highlighted as the chosen view for this research.

Definition of the term 'Epistemology'. The term *Epistemology*[5] has suffered from several semantic changes. Nowadays, Epistemology is the branch of Philosophy that studies knowledge. There is no consensus for a commonly shared definition. Going through the most common dictionaries, one finds the following interpretations: the *Oxford Dictionary* defines it as "[t]he theory of knowledge, especially with regard to its methods, validity, and scope, and the distinction between justified belief and opinion."[6] While the *Merriam-Webster Dictionary* says that it is "the study or a theory of the nature and grounds of knowledge especially with reference to its limits and validity".[7] Turning to specialised literature, Hendricks says that "[t]he systematic and detailed study of knowledge, its criteria of acquisition and limits and models of justification is known as epistemology" (Hendricks 2005, p. 1).

Rivers of ink have already flown regarding the different interpretations of the term and in which sense it is related to all types of knowledge or information. For the present purpose I will stick to the first and simple definition I gave above: *branch of Philosophy that studies knowledge*. As such, there are many different aspects regarding the notion of knowledge that may fall under the epistemological investigations. Of course not every approach to knowledge is suitable of being considered Epistemology, but nowadays this field of study has widen its range in such a way, that aspects like perception, communication, social factors or cognitive processes are also part of it.

Origins of Epistemology. The topic of Epistemology has been one of the main themes in philosophical discussion since its very beginning in Ancient Greece. One could establish its origin in Plato's dialogues. It is said that Plato states that knowledge is equivalent to *justified true belief* (JTB) in the *Meno* and the *Theatetus*. Though this claim is not accurate and those who state it in these terms are accused of anachronism,

[5] Despite of the fact that the topic is very old, the term itself was coined for the first time in Ferrier (1854).

[6] In https://www.lexico.com/definition/epistemology, accessed 01/02/2020.

[7] In https://www.merriam-webster.com/dictionary/epistemology, accessed 01/02/2020.

what may be claimed without falsehood is that in these dialogues, Plato equals 'knowledge' with 'true belief' and adds some additional condition, which might be called 'justification' in modern terms. Attributing modern terms to ancient savants is always complicated and the lack of accuracy is a constant thread. Thus, it should suffice to claim that the origin of Epistemology can be traced back to Plato's dialogues where the thesis that 'knowledge is justified true belief' was originated.

The JTB-view (as Plato's understanding is called) is often used to define traditional Epistemology by itself. The different ways in which 'justification' can be interpreted or understood will give rise to different points of view or branches of Epistemology. The most general classification is made regarding the source of the justification. Thus, on the one hand, there is *Epistemic Internalism*, that claims that the justification for true beliefs happens 'inside' the subject, it is an internal matter where the actual world plays no role at all; and, on the other hand, one finds *Epistemic Externalism*, stating that whatever justifies a true belief must stem from 'outside' the subject, that is, it is the external world that determines the justification. Both, Epistemic Internalism and Externalism, are still being continued in mainstream Epistemology and promote endless arguments and counterarguments, as will be explained below.

Going back to the historical foundations of Epistemology, specifically Epistemic Internalism, the JTB-view was 'adopted', many centuries later, by René Descartes. I risk here being accused guilty of anachronism, since Descartes himself never uses 'justified true belief'. What is clear, is that in Descartes' Rationalism knowledge is all that of which no subject can doubt about its truth and is achieved by means of reasoning. And since *all that* might be everything the subject believes, one could somehow claim that he refers to *justified true belief*. Descartes' Epistemology is seen as one of the first variants of Internalism. The distinction between Internalism and Externalism goes hand in hand with the difference between Rationalism and Empiricism, respectively, contemporary to Descartes' era.

Currently, the view by René Descartes is referred to as *the Cartesian view* and it is assumed to suffer from two big epistemic problems: *scepticism* and *solipsism*.

Scepticism, knowledge is problematic or doubtful, represents Descartes' biggest challenge from the very beginning of his philosophy.[8] As such, Epistemic Internalism is suitable of leading to scepticism. The possibility of knowledge is questioned here and it is quite complicated to find a trustful and truthful source for knowledge in the Cartesian view, since his metaphysical convictions forced him to doubt absolutely every source of knowledge. As a consequence of that, Descartes' theory also falls into *solipsism*, the self is the only thing to know. The ultimate consequence of the Cartesian method, after doubting every source of truth is that only the self can be known, the problem arises when he tries to build new knowledge stemming from the self and doubting every external source. The ways in which Descartes solves this problem and the suitability of it is of no concern to this research, hence I will just close this brief historical review by mentioning that Descartes appeals to God as source of omniscience to help the solipsist subject to gain knowledge again.

[8] See, for instance, Chamizo-Domínguez (1984) and Gilson (1950) for a discussion about the origin of Descartes' thoughts and motivations.

Currently, one could say that the Cartesian view of knowledge established that "it is a conceptual truth, that if conditions C justify belief B for subject S, then C logically entails that B is true" (Cohen 1984, p. 280). Being this so, the step to scepticism is straightforward: any conditions C may be established, that make B false.

Epistemic Thought Experiments. One of the most common argumentative course of action used in mainstream Epistemology is based on thought experiments. An epistemic thought experiment being a theoretical and ideal situation where a subject (or more) is put in an idealized, and sometimes implausible, context, where she experiences some type of delusions or misconceptions. Most thought experiments are created for arguing in favour of either Internalism or Externalism.

Probably, one of the most well known thought experiments that was coined in the past is the *Evil Demon Case*, suggested by Descartes, in order to achieve the universal doubt.[9] In contemporary epistemological approaches there has been a remake of this experiment, denoted as the *New Evil Demon Case*.[10] This experiment, similar to the original, consists in a situation where there is an Evil Demon fooling the subject in a sense that every information she obtains by perception (empirical knowledge) seems false, though it is true. With the New Evil Demon Case the internalists prove that the external factors do not play a role in the acquisition of knowledge and, even more, that paying attention to them could lead to false beliefs and misconceptions.

Although it is not exactly a thought experiment, I have to mention at this point the famous *Gettier-Cases*. In Gettier (1963), a paper that is only three pages long, the author states one of the most important challenges to Epistemic Internalism. Gettier presents two plausible cases that show that a subject can have false beliefs that, nevertheless, serve as justifications for true propositions; resulting in scenarios where a subject 'knows' a proposition that is true, but her justifications for it are based on false beliefs that accidentally turned out to validate the true proposition.

On the other side, regarding experiments made up by externalists to prove their point, one could highlight the *Twin Earth Thought Experiment*, presented in Putnam (1973). This experiment is based on the idea that there is a Twin-Earth, where everything looks the same as on Earth, but the chemical composition of the water differs from the water on Earth, hence twin-water is made of XYZ instead of H_2O. Then, he presents a subject, Oscar, that lives on Earth, and Twin-Oscar, that lives on Twin-Earth; and proposes a scenario where the subjects switch worlds, such that now Oscar is on Twin-Earth and Twin-Oscar is on Earth. Both are swimming in a pool and they know that they are swimming in water, but their thoughts will be false, since Oscar thinks he is swimming in H_2O, while swimming in XYZ and vice versa. What this thought experiment tries to prove is that the external world determines the semantic content of our thoughts and hence also the justifications of them.

[9]Though the term 'thought experiment' corresponds to the common terminology of Epistemology, some specialists on Descartes would claim that this example is a hypotheses, instead of an experiment.

[10]Stated for the first time in these terms in Cohen (1984).

As should be already clear, neither of those thought experiments are absolute or definitive when it comes to proving one's point. The argumentative style adopted in mainstream Epistemology has resulted in a very long list of literature that not only discusses and rewrites these (and other) thought experiments, but also creates new versions of both Internalism and Externalism.

Contemporary Epistemology. Contemporary Epistemology is such a wide-ranged field that one needs to delimit it, if the term 'Epistemology' is intended to denote something at all. To do so, I will allude to Hendricks, who draws the difference between 'Mainstream and Formal Epistemology' in his homonymous book in the following sense:

> Contemporary epistemological studies are roughly carried out: (1) in a *mainstream* or informal way, using largely conceptual analyses and concentrating on sometimes folksy and sometimes exorbitantly speculative examples or counterexamples, or (2) in a *formal* way, by applying a variety of tools and methods from logic, computability theory or probability theory to the theory of knowledge. The two traditions have unfortunately proceeded largely in isolation from one another (Hendricks 2005, p. ix. Original italics).

'Formal Epistemology' corresponds, in a wide sense, to Epistemic Logic and its developments. As explained in the previous section, EL is a logical representation of knowledge (and doxastic logic, of belief) but it has little to do with mainstream Epistemology. As Hendricks himself points out: "it is a discipline devoted to the logic of knowledge and belief, but is alien to epistemologists and philosophers interested in the theory of knowledge" (Hendricks 2005, p. 81). Therefore, in what follows, I will continue this exposition sticking to mainstream Epistemology and turn now to contemporary approaches.

To reassure what is meant here with 'mainstream Epistemology' and erase every whisper of doubt there might still be, I will quote again Hendricks, who, in a more revealing way, states the following about how mainstream Epistemology is conducted:

> The term 'mainstream epistemology' refers to the *modus operandi* of seeking necessary and sufficient conditions for the possession of knowledge based on the standard definition or some close derivative thereof. [...] It is a dialectical and sometimes even 'diabolical' process that by its very nature balances between the theory of knowledge and skepticism (Hendricks 2005, pp. 14–15. Original italics).

Fortunately, in contemporary Epistemology, there are many authors who claim to have solved this 'scepticism-alert', at least from an internalist point of view. Epistemic Internalism stems from the aforementioned *Cartesian View*, defined by Cohen in terms of the 'conditions C' that justify 'belief B' for 'subject S', such that 'C logically entails B'. This view results very easily in a vicious circle, when C is compatible with B being false; or when a *Gettier-Case* is explained in those terms.

Most forms of contemporary Epistemic Internalism share, thus, a common ground: they impose some *additional restrictions* on conditions C such that the vicious circle can be broken and the true belief is guaranteed.

These restrictions deal with the way in which the notion of *justification* is interpreted and how the subject accesses the information at her disposal. Of course the label 'internalism' is related to an internal source for knowledge; but this does not conflict with the claim that the *grounds for a justification* may come from an external source. Some epistemologists have claimed, that accepting an external source leads directly to some form of Externalism, I prefer to consider it a more intuitive and accurate understanding of the complex process of gaining knowledge, where the creation of knowledge is, indeed, internal to the subject, but the origin of the information known by her, stems from her perception of the world (external source). Hence, Internalism and Externalism, views that started out being complete opposites, represent two elaborated and wide-ranged theoretical frameworks that may share common positions without colliding.

Contemporary Epistemic Internalism. The frontiers between Internalism and Externalism are being blurred in increasingly more approaches. Nevertheless, for the purpose of this research, I am interested in the new approaches to Epistemic Internalism. Thus, I will briefly sum up the contemporary perspectives that fall under the designation of 'Internalism'.

In Pappas (2017) one finds a very extensive review of the different understandings of 'epistemic justification' from both the internalist and externalist point of view. What I want to highlight from this classification are the different useful concepts when contemplating the overall picture. First, the author names *Knowledge Internalism* all those approaches that deal with "knowing or being aware of that on the basis of which one knows" (Pappas 2017, para. 1), as analogous to what justifies the beliefs.

In *Knowledge Internalism*, there are different interpretations. What is common to all of them is the importance that the concept of *accessibility* has. How and to which amount the subject accesses the *knowledge basis* that acts as a justifier will determine the type of *Accessibility or Awareness Knowledge Internalism* that is being proposed. In this classification there is a *strong* and a *weak* version of Awareness Knowledge Internalism, with the former forcing the subject to be aware of something that is indeed in her knowledge basis, and the latter demanding the subject to be aware of everything that is in her knowledge basis.

Since, as explained at the beginning of this section, 'knowledge' is sometimes equated with 'justified belief', there are versions of these same proposals that instead of focusing on 'knowledge', are centred on 'justification'. In this manner, one finds a weak and a strong version of *Awareness Justification Internalism* (and also Externalism, that corresponds to the denial of the internalist proposal). The main difference between the 'Knowledge Internalism' and 'Justification Internalism' being that instead of a 'knowledge basis' that justifies the belief, the latter proposes a *justifier* as the umbrella term for all that information that is suitable of justifying other beliefs that will become knowledge.

While accuracy in the terms is always wanted, I will stop making detours and take both views as analogous accounts for the *Awareness Justification Internalism*. This view is one of the contemporary proposals that overcomes the problem of scepticism,

their terminology. But is it really like this? Well, in the big picture it certainly is, and that is why the disconnection is so clearly perceived. But as will be shown below, there is still much hope for a new convergence.

Tending bridges During the last two decades, there have been some attempts to bring both fields back together, papers like Hendricks and Symons (2005) and van Benthem (2006) are good representatives of those attempts. The main goal of these papers, and also collections such as Arló-Costa et al. (2016), is to unravel the interconnections that have always been there, but are sometimes difficult to observe.

Hendricks and Symons find the 'bridges' they are looking for in the central notions that define both fields, namely "knowledge, belief and doubt" for mainstream Epistemology, and "learning, information and strategies" for Epistemic Logics. They state that "these two sets of notions are congruent and parallel" (Hendricks and Symons 2005, p. 160). What these authors show in their paper is that though with different technicalities, Epistemology and Epistemic Logic are still dealing with the same problems and solving them in quite similar ways. Their general attitude is based on the idea, that there are several notions that act as a bridge between both fields and that future proposals will learn from both.

Turning now to van Benthem (2006), the author shows a positive view regarding the interconnections of Epistemology and Epistemic Logic. As mentioned before, there is a thread of hope that both areas will merge again when applying the dynamic perspective to information, and this is exactly what van Benthem does in this paper. He reviews most of the central notions from the new paradigm of the studies of language and information (which, in turn, determine the new Epistemic Logics that are being developed today) and concludes that these concepts are not only present in mainstream Epistemology, but some may also stem from there. He is also positive about the future interactions and, to sum up, presents what I like to call a 'virtuous circle', where the results of a collaboration can only be those of a mutual benefit.

Parallel development of Epistemology and Epistemic Logic. It is clear from the foregoing, that Epistemology and Epistemic Logic represent two different research areas with its own terminology and methods, and that there is a common historical background to both fields that will help me bring them back together. For doing so, it is very helpful to present the development of both fields in parallel, showing thus the interconnections this book will analyse.

Regarding the tradition, or historical origin, Epistemology stems from the dispute between Internalism and Externalism regarding the justification of the 'true belief', while Epistemic Logic has its origin in the extension of Modal Logic that applies its structure to the knowledge of epistemic agents. Both subjects have been challenged with a problem that threatened their very existence. Epistemology, specifically Epistemic Internalism, suffered from the problem of Scepticism, which for years has been the burden every internalist had to solve. On the other hand, Epistemic Logic was challenged with the problem of Logical Omniscience.

The respective solutions to these challenges share the same structure, namely to impose some kind of restriction to the main concept that is being questioned. In the

case of Epistemic Internalism it is the concept of *justification* that will require some additional restrictions, while in Epistemic Logic it will be the *knowledge* (in most cases the explicit version) that has to be restricted due to some additional element.

As already mentioned regarding both areas, there are many different solutions, but the two I will consider along this research are the view of Awareness Justification Internalism (AJI) for Epistemology and Awareness Logic (AL) for Epistemic Logic. To continue this parallel presentation the concept that is added to both fields as the new element that solves the respective problem is the concept of *awareness*. In AJI in the form of the Awareness Requirement, that limits the justification, and in AL as the awareness operator, that serves as a filter on knowledge. I state that for the purpose of the present research the concept of awareness is the appropriate tool. To which extend this is the case will be revealed in what follows.

The concept of Epistemic Awareness. So far, every appearance of the term 'awareness' has been made in its 'epistemic' understanding, but before continuing an argument in favour of this concept, I think that a more thorough specification of the notion is needed. 'Awareness' is a very polysemic term, sometimes equated with 'consciousness' and others applied to a long list of perceptual features of human beings. This ambiguity brings me to the need of clarification.

I connect 'awareness' with an epistemic understanding of 'perception' (the information one perceives)[11], and 'consciousness' with all types of perceptions, including their moral and psychological implications. That is, the thoughts about 'reasons for actions' or 'the right thing to do' may be part of what 'consciousness' alludes to in a general picture, while they should never be part of a subject's awareness. Of course, in more informal contexts both terms are used as synonymous and it is common to read phrases like 'raise of awareness' or 'being conscious that something is the case', but with the former distinction and the definitions below I intend to overcome this confusions and establish a common terminological background for this research.

The 'awareness' I will be referring to is always epistemic, that is, it accounts for the fact of *realizing one's own information*. Attending the standard definitions, we find in the *Oxford Dictionary*[12] 'awareness' as "[k]nowledge or perception of a situation or fact", and in the *Merriam-Webster Dictionary*,[13] "the quality or state of being aware: knowledge and understanding that something is happening or exists."

Though neither of these definitions allude to the verb 'to realize' I mentioned in my own definition, it is somehow implied, since the definitions of this verb include the notion of 'awareness'. The first entry of 'to realize' in the *Oxford Dictionary*[14] says "[b]ecome fully aware of (something) as a fact; understand clearly"; and the third entry of the *Merriam-Webster Dictionary*[15] defines it as "to conceive vividly as real: be fully aware of".

---

[11] In many places, like MacMillan (2012), this understanding is referred to as 'conceptual awareness'.

[12] In https://www.lexico.com/definition/awareness, accessed 01/12/2020.

[13] In https://www.merriam-webster.com/dictionary/awareness, accessed 01/12/2020.

[14] In https://www.lexico.com/definition/realize, accessed 01/12/2020.

[15] In https://www.merriam-webster.com/dictionary/realize, accessed 01/12/2020.