

## **Robustness and Reality**

*Synthese*, DOI: 10.1007/s11229-015-0801-6

Markus I. Eronen

KU Leuven

Institute of Philosophy

Centre for Logic and Analytical Philosophy

Kardinaal Mercierplein 2 - box 3200

3000 Leuven

markus.eronen@hiw.kuleuven.be

### **Abstract**

Robustness is often presented as a guideline for distinguishing the true or real from mere appearances or artifacts. Most of recent discussions of robustness have focused on the kind of derivational robustness analysis introduced by Levins, while the related but distinct idea of robustness as multiple accessibility, defended by Wimsatt, has received less attention. In this paper, I argue that the latter kind of robustness, when properly understood, can provide justification for ontological commitments. The idea is that we are justified in believing that things studied by science are real insofar as we have robust evidence for them. I develop and analyze this idea in detail, and based on concrete examples show that it plays an important role in science. Finally, I demonstrate how robustness can be used to clarify the debate on scientific realism and to formulate new arguments.

## 1. Introduction

Robustness refers to multiple derivability, detectability or measurability. This idea may go all the way back to Aristotle, but the author who is usually credited for introducing robustness to the contemporary debate is Levins (1966), who argued that we can use several models with different simplifying assumptions to derive a robust result. Wimsatt (1981) embraced this idea, while authors such as Orzack and Sober (1993) strongly criticized it. In recent years, there has been a resurgence of work on robustness (e.g., Weisberg 2006, Woodward 2006; Kuorikoski, Lehtinen & Marchionni 2010; 2012, Soler et al. 2012, Raerinne 2013, Hudson 2014).

A tendentious idea that is associated with robustness is that it may provide justification for the step from “mere models” to what is actually real. This idea comes up in different forms in the work of the proponents of robustness. Levins (1966) famously asserted that the truth is at the intersection of independent lies. According to Wimsatt (1981, 1994), robustness is a criterion for the real. Weisberg (2006) and Kuorikoski et al. (2010, 2012) argue that robust theorems characterize or give information about the underlying causal relationships.

As many authors have pointed out, there are in fact several different forms of robustness that only share the very general basic principle. Most of the recent discussion has focused on the sort of robustness analysis introduced by Levins and elaborated by Weisberg (2006), Woodward (2006), Kuorikoski et al. (2010, 2012), and others (from now on derivational robustness, or DR). The idea of DR is that a modeler can use several idealized or simplified models of a system or phenomenon to derive a robust result that does not depend on any of the individual simplifying assumptions. There is an ongoing debate concerning the extent to which robustness analysis can increase our confidence in the truth of the robust theorems, or the reality of the causal mechanisms involved. However, a more general form of robustness that is potentially more relevant for justifying inferences to what is real has received less attention. This is the kind of robustness that Wimsatt (1981, 1994) is mainly concerned with: things are robust if they are detectable, derivable, producible, etc., in

a variety of independent ways (robustness as multiple accessibility, MA). In the literature, these two forms of robustness are often conflated, but as I will briefly show, they differ in such crucial ways that it is better to keep them clearly distinct.

In the main part of this paper, I will defend the idea that Wimsatt's notion of robustness as multiple accessibility (MA), when properly understood and defined, can provide justification for our ontological commitments. What this means is that we are justified in holding something to be real if it is robust, that is, if it is detectable, producible, or derivable in a variety of independent ways. I will elaborate and refine this view, and address several possible problems. I will also illustrate the approach with examples from science, and finally consider the implications it has for the issue of scientific realism.

As I will argue, robustness understood as multiple accessibility (MA) fulfills an important and neglected role in contemporary philosophy. Philosophers have traditionally focused on searching for a metaphysical criterion or definition for what is real, which results in very general claims such as "all real things are composed of fundamental physical particles" (e.g., Pettit 1993). Such general principles concern the final and absolute make-up of the world, and may be important in purely metaphysical projects, but are of little use if we are concerned with the (fallible) ontological commitments that current science justifies. In other words, traditional metaphysical approaches provide little help in answering the question: What are we justified in holding to be real now, as limited beings, based on the current state of science? As I will show in this paper, robustness (MA) provides one answer to this question.

## 2. Robustness

Richard Levins (1966) introduced the concept of robustness to the debate on modeling and explanation. He observed that the systems studied in population biology are so immensely complex that it is practically impossible to represent them

faithfully in all their detail, so in their models, scientists need to make simplifying assumptions or idealizations, as well as tradeoffs between values such as precision and generality. This opens the unfortunate possibility that the results obtained may be just artifacts of the simplifying assumptions. However, if the modeler can apply several independent models of the same system with different simplifying assumptions and they all yield a similar result, she can be more confident that the result depends on the essential parts of the models and not on the simplifying assumptions. Such a result is “a robust theorem”. Robust theorems are typically predictions or generalizations, such as “in an uncertain environment, species will evolve broad niches and tend toward polymorphism” (Levins 1966, 423). Since models with strictly speaking false assumptions can result in robust theorems, “our truth is the intersection of independent lies” (ibid.). Although Levins’ topic was population biology, this idea is entirely general and can be applied to other contexts where idealized models are used, which includes most fields of science.

Wimsatt (1981) discussed and expanded Levins’ ideas from a broader philosophical perspective, and recently authors such as Weisberg (2006; Weisberg & Reisman 2008) and Kuorikoski et al. (2010, 2012) have continued this project. For example, building on the work of Weisberg, Kuorikoski et al. (2010) argue that modeling in economics can be seen as (collective) derivational robustness analysis. In derivational robustness analysis, modelers make substantial assumptions about the system or phenomenon under study, and then use robustness analysis to derive a common result from different models, each including the same substantial assumptions but different simplifying or “tractability” assumptions. The substantial assumptions are hoped to be realistic, while the tractability assumptions are known to be (strictly speaking) false. The common result, a robust theorem, gives information about the causal relationships in the system under study. The illustration given by Kuorikoski et al. (2010) is modeling in geographical economics. In this case the tractability assumptions include things such as specific utility functions or unrealistic but straightforward functions of transportation costs (ibid., 556-557). From various models with different tractability assumptions, the following robust

theorem can be derived: “Ceteris paribus, spatial agglomeration occurs when economies of scale are high, market power is strong, and transportation costs are low” (ibid., 555).

This derivational robustness analysis (DR) has been criticized from different angles (Orzack & Sober 1993, Woodward 2006, Forber 2010, Odenbaugh & Alexandrova 2011). In particular, the explanatory or confirmatory role that DR plays in science has been a matter of much debate in recent years. There is also no agreement on the relation between robust theorems and the underlying causal mechanisms. However, all of the authors involved in the debate agree that DR *alone* does not confirm the existence of causal relationships, or justify inferences to what is real. Some kind of empirical confirmation is always needed, be it then low-level confirmation of the mathematical framework (Weisberg 2006), confirmation of the substantial assumptions involved (Kuorikoski et al. 2012) or confirmation of the robust theorem itself (Odenbaugh and Alexandrova 2011). In sum, DR certainly plays an important role in science, but its relation to ontological commitments and the extent to which it allows inferences to truth or reality is far from clear. Since the issue has been extensively debated by the authors mentioned above, I will not discuss it in more detail here.

While DR has been thoroughly analyzed in recent literature, Wimsatt’s (1981/2007, 1994) work on robustness has received relatively little attention (notable recent exceptions being Soler et al. (2012) and Hudson (2014)). Wimsatt argues that the general idea of robustness has been at the background throughout the history of science, and can be found in the work of authors such as Aristotle, Galileo, Whewell, and Peirce, but is rarely made explicit and usually appears in footnotes or parentheses (1981/2007, 43).<sup>1</sup> Wimsatt characterizes robustness analysis as consisting of the following procedures: “1. To analyze a *variety of independent* derivation,

---

<sup>1</sup> An anonymous referee also pointed out that Husserl discusses criteria resembling robustness. A thorough historical study would probably reveal many further sources for the idea.

identification, or measurement processes. 2. To look for and analyze things that are *invariant* over or *identical* in the conclusions or results of these processes. 3. To determine the *scope* of the processes across which they are invariant and the *conditions* on which their invariance depends. 4. To analyze and explain any relevant *failures of invariance*.” (Wimsatt 1981/2007, 44). Things that are invariant under such an analysis are *robust* (ibid.). According to Wimsatt (1981/2007, 45-46), the “processes” in clause 1 may refer to a broad range of things: different sensory modalities, experimental procedures, assumptions, models, axiomatizations, tests, scales, different sets of microstate conditions, theoretical descriptions of the same phenomenon at different levels of organization, and so on.

Importantly, Wimsatt also argues that robustness is the criterion for distinguishing real things from mere appearances or artifacts: “[A]ll the variants and uses of robustness have a common theme in the distinguishing of the real from the illusory; the reliable from the unreliable; the objective from the subjective; the object of focus from artifacts of perspective; and, in general, that which is regarded as ontologically and epistemologically trustworthy and valuable from that which is unreliable, ungeneralizable, worthless, and fleeting” (ibid., 46). The reason why robust things are reliable and ontologically/epistemologically trustworthy is that if there are many independent ways of deriving or measuring something, the probability that all those ways happen to be mistaken is the product of the individual probabilities of error, and thus declines very rapidly as new independent ways are added (ibid., 47-50).<sup>2</sup> This general idea comes up in different forms also in Chang (2004), Franklin & Howson (1984), Hacking (1983), Salmon (1984), Staley (2004) and Trout (1998), among many others (see Stegenga 2009 for more).

In a later article, Wimsatt (1994/2007) continues developing this approach, and explicitly presents robustness as a criterion for the real: “Although nothing will

---

<sup>2</sup> This central robustness argument seems to be an instance of an inference to the best explanation, and resembles the “no miracles” argument for scientific realism (see Hudson 2014 for more on how these three are related). I return to the issue of robustness and scientific realism in section 5 below.

guarantee freedom from error, robustness has the right kind of properties as a criterion for the real” (Wimsatt 1994/2007, 197). He gives the following characterization of robustness: “*Things are robust if they are accessible (detectable, measurable, derivable, definable, producible, or the like) in a variety of independent ways*” (ibid., 196). The expression “things” is to be taken very generally, and potential targets of robustness include entities, properties, relations, propositions, and levels (ibid.). This idea of robustness as multiple accessibility (MA) is what I will focus on in the rest of this paper.<sup>3</sup>

Wimsatt tends to see robustness as a general principle that is constantly working in the background in science. In this spirit, one could argue that derivational robustness (DR) is just a special case of MA, such that the target is a robust theorem, and the independent ways of deriving it are models with different simplifying assumptions (cf. Kuorikoski et al. 2010, 544-545, Wimsatt 1981/2007, 56-57). However, DR does not involve any independent ways of producing, measuring, or detecting the phenomenon, over and above applying different models. It is questionable whether it makes sense to think of different models of the same system (and with the same substantial assumptions) as sufficiently “independent ways” to yield MA (Odenbaugh & Alexandrova 2011, 762-763). Moreover, the results of DR are theorems (propositions), while MA primarily concerns entities, properties and phenomena – although robust theorems in DR refer to entities and properties, what is robust in DR is the theorem itself, and not the entities or properties.<sup>4</sup> MA is presented as criterion for the real, but none of the proponents of DR argue that it should be seen as such a criterion. DR is closely linked to the idea that a modeler typically has to make

---

<sup>3</sup> The distinction between DR and MA was clearly and succinctly pointed out by Calcott (2011), who called the former “robust theorems” and the latter “robust detection”.

<sup>4</sup> One could argue that also the results of MA are in the end propositions, for example propositions stating that some entity exists or has some property. However, even with this interpretation, there is a difference in kind between DR and MA: the results of the former are propositions of one kind (theorems), while the results of the latter are propositions of a different kind (existential propositions or property attributions).

simplifying assumptions and cannot maximize all epistemic values (such as precision and generality) simultaneously, resulting in tradeoffs between different values (Levins 1966, Matthewson & Weisberg 2009), while for MA this idea of tradeoffs and simplifying assumptions is inessential. In general, the context where DR is applied is modeling, while MA is a broader metascientific principle that plays an important background role in science (see sections 4 and 5).<sup>5</sup> For these reasons, we should not consider DR to be a special case of MA, but rather a different form of robustness (see also Calcott 2011, who reaches a similar conclusion).

Importantly, the classic criticism raised against robustness concerns DR, and does not apply to MA. The main point of this criticism is that DR constitutes a problematic non-empirical form of confirmation, since one can derive a robust theorem from models without doing any experiments or observations (Orzack & Sober 1993, Odenbaugh & Alexandrova 2011). The supporters of DR have argued that this is not the case, appealing for example to the prior empirical confirmation of the (substantial assumptions) of the models involved (Kuorikoski et al 2012). However, MA typically involves independent ways of *empirically* detecting or measuring the target phenomenon, and thus escapes this criticism altogether.<sup>6</sup> Of course, there are various other potential problems involved with MA; I will return to them in Section 3.

It should be noted that there are various other notions or categories of robustness in the literature (cf. Boon 2012, Calcott 2011, Raerinne 2013, Weisberg & Reisman 2008, Woodward 2006). These forms play different roles in science, and warrant more detailed analyses, but this is beyond the scope of this paper. I have taken DR as a

---

<sup>5</sup> Results from modeling can also contribute to MA by indicating that the target entity or property plays an explanatory role, but this is neither necessary nor sufficient for MA – see section 3.

<sup>6</sup> One implication of my account presented below is that there is a (low) degree of robustness also in the special case where there are just several independent ways of *deriving* a phenomenon. In this case, the criticism of non-empirical confirmation may apply. However, this is just a limit case, and nothing important turns on it: if empirical confirmation is taken to be necessary, we can simply add a clause stating that independent ways of derivation alone are not sufficient for robustness.



starting point since it has been the focus of the recent philosophical discussion on robustness (Kuorikoski et al. 2010, 2012, Odenbaugh & Alexandrova 2011, Weisberg 2006, Weisberg & Reisman 2008), and in the rest of this paper I will concentrate on robustness as multiple accessibility (MA) and its potential role as a criterion for the real.

### 3. Robustness as a criterion for the real

The idea that robustness as multiple accessibility (MA, from now on simply robustness) is a criterion for the real immediately raises the need for clarification. Several features of robustness make it *prima facie* problematic as a criterion for the real. First of all, robustness is a matter of degree: some things are more robust than others. For example, there are numerous independent ways of measuring, detecting, or deriving the temperature of an object, but (as of yet) very few independent ways of measuring, detecting, or deriving the Higgs boson. In contrast, taking reality to be a matter of degree is problematic: the idea that some things are more real than others is something many (though not all) philosophers would reject out of hand. Furthermore, our conceptions of how robust things are may change over time – there are far more ways of measuring, producing, or deriving electrons now than there were in the times of Thomson and Rutherford. There is also no guarantee that things that are considered robust based on the current state of science are real, as arguments such as pessimistic induction purport to show (e.g., Laudan 1981). Robustness is clearly a fallible criterion that is relative to a certain scientific community at a certain time.

Due to these features, we should take robustness to be an *epistemic* criterion : a criterion that tells us when we have good grounds to consider something to be real. This also seems to be what Wimsatt primarily has in mind: “I want criteria for what is real which are decidedly local – which are the kinds of criteria used by working scientists in deciding whether results are ‘real’ or artifactual, trustworthy or

untrustworthy, ‘objective’ or ‘subjective’” (Wimsatt 1994/2007, 195). Unfortunately, Wimsatt never clearly explains what this exactly means and how robustness is supposed to function as an epistemic criterion.

I propose to clarify this idea and its epistemic nature by spelling it out in terms of justification: roughly, we are *justified* in believing that X is real insofar as X is robust given the state-of-the-art scientific research (see also Eronen 2012). With this approach, the features of robustness characterized above are easily captured. Since robustness is a source of justification, it is clear that robustness can be a matter of degree, and that conceptions of what is robust can change with time and the scientific community: we are considerably more strongly justified in believing that H<sub>2</sub>O molecules are real than we are in believing that the Higgs boson is real, and we have a far stronger justification for holding electrons to be real nowadays than in the late 19<sup>th</sup> century.

However, before I can formulate this idea more precisely, the definition of robustness itself needs to be updated. As we saw in the previous section, Wimsatt gives the following rough definition of robustness: *Things are robust if they are accessible (detectable, measurable, derivable, definable, producible, or the like) in a variety of independent ways.* This definition is problematic for several reasons. First of all, it does not incorporate the idea (mentioned above) that robustness is a matter of degree. Secondly, it defines robustness entirely generally, and does not make it relative to a scientific community at a certain time.<sup>7</sup> Third, the statement that robust things need to be “accessible” in different ways is vague, and the expression “... or the like” in the parentheses leaves the different ways of access that can contribute to robustness entirely open. Furthermore, it is not clear whether the different types of ways of access (detection, measuring, derivation etc.) should be seen as necessary or sufficient conditions for robustness, or how exactly the ways need to be independent from each other.

---

<sup>7</sup> Wimsatt also mentions these features of robustness; my point is merely that they are not reflected in the rough definition he gives.

I will proceed here by first giving an updated definition of robustness. I then explain and clarify the different parts of this definition, and after that turn to the variety and independence of the ways of access. The updated definition is as follows:

(Robustness) X is robust in the relevant scientific community at a certain time insofar as X is detectable, measurable, derivable, producible, or explanatory in a variety of independent ways.<sup>8</sup>

In this revised definition, the target of robustness is intentionally left open (“X”). I take the definition to apply primarily to entities, properties, and phenomena, but do not rule out the possibility that it might apply to other sorts of things as well. As we saw in section 2, Wimsatt takes his rough definition to apply also to relations, propositions, levels, etc., but since extending robustness to these categories might bring along additional problems, I refrain from doing it here, and focus on the robustness of entities, properties, and phenomena.

In the definition, the expression “insofar as” is intended to capture the fact that robustness comes in degrees. The vague reference to different ways of access has been replaced with an explicit list. All of these ways can contribute to robustness, but none of them should be seen as a necessary condition (a point to which I return below).

What I mean by “explanatory” in the new definition is roughly that the candidate for robustness appears in explanatory generalizations (for example, invariant generalizations as in Woodward (2003)) or models of the relevant scientific community. I have added this to accommodate the widely held view that the fact that a kind appears in scientific laws or explanatory generalizations is (fallible) evidence

---

<sup>8</sup> Strictly speaking, it would be more exact to say that “evidence for X is robust” or “our access to X is robust” instead of “X is robust”, since the latter expression seems to already imply the existence of X, and rule out by definition the robustness of non-existent entities (such as phlogiston). However, in most contexts nothing important turns on this distinction, so for the sake of readability, I continue to talk of X itself as robust (instead of our evidence for it).

for its reality (e.g., Bird 1998, 74, Boyd 1991, 139-140). The idea is that if an entity or property does not appear in any explanatory generalizations or models, it has a relatively low degree of robustness, while an entity or property (such as the electron) that appears in a broad range of different explanatory generalization or models has a relatively high degree of robustness.

I have also removed “definable”. In Wimsatt’s original context it is justifiable, since Wimsatt intends robustness to apply also to concepts, laws, and propositions (Wimsatt 1981/2007, 55). Robustness through multiple definability seems suitable (or at least potentially interesting) when applied to linguistic, theoretical, or mathematical objects. However, it is far from clear how (independent) definability could contribute to the robustness of entities, properties, or phenomena. There are various ways of defining what a gene is, but does that make the property of being a gene more robust? Clearly not, since the different definitions result in different categories with different extensions. Since the main focus here is on the robustness of entities, properties, and phenomena, I have not included definability in the definition.<sup>9</sup>

The revised definition also reflects the fact that the different dimensions that can contribute to robustness (detectability, measurability, etc.) are not necessary conditions for robustness. This is an important point that has not been made explicit enough in the literature.<sup>10</sup> For example, a phenomenon that is detectable, measurable,

---

<sup>9</sup> It is also likely that what Wimsatt means by ‘definition’ or ‘definability’ is something weaker than what philosophers usually mean by it. For example, a definition of an electron given by a scientist could be that the electron is an elementary particle that has a negative electric charge of about  $1.602 \times 10^{-19}$  coulombs and a mass of about  $9.11 \times 10^{-28}$  g. A philosopher would see this as a characterization of an electron, and not as a definition. In the context of robustness, multiple definability makes much more sense if we understand it in the weaker sense, i.e., as multiple characterizability. I thank Paul Teller for pointing this out.

<sup>10</sup> To rule out a further potential source of misinterpretation: detectability, measurability, etc., does not refer to detectability, measurability, etc., *in principle*, or to physical or metaphysical possibility. It refers to detectability, measurability, etc. with current technology, experimental setups, theories, and so on.

and derivable in a variety of independent ways, but is impossible to produce due to technological or practical (or even ethical) limitations, can be highly robust. Examples of this kind of phenomena include astronomical phenomena such as solar flares and supernovae. On the other hand, a phenomenon that is detectable, measurable, and producible in a variety of independent ways, but cannot be explained or derived from any existing theories, can be highly robust. For example, the spectrum of black-body radiation was such a robust phenomenon, until Planck showed how it can also be derived from theory. In fact, even one dimension may be sufficient for a (low) degree of robustness: If there are several independent ways of detecting a phenomenon, it has a degree of robustness, even though it cannot be produced, measured or explained. Examples of this kind include unexplained astronomical phenomena, such as supernovae observed before the 20<sup>th</sup> century. What is crucial is that the ways are sufficiently independent (see below), not that they are of different types. However, it is plausible that adding different types of access (measurement, detection, production, etc.) increases robustness more than adding more ways of the same type.<sup>11</sup>

The definition of robustness refers to *independent* ways of detecting, measuring, etc. This is related to the principle mentioned in section 2: If we come up with more and more ways of accessing something, the probability that all those ways happen to be

---

<sup>11</sup> Naturally, the different ways of detecting, measuring, or producing a phenomenon can also vary regarding their inherent relevance: some methods are more reliable and produce stronger evidence than others. However, the issue of evaluating strength of evidence has been discussed in other contexts in philosophy of science (see for example Crupi, Tentori & Gonzales 2007 or Cartwright 2007, Ch. 3 for more). Since the strength of different lines of evidence can be evaluated independently of considerations of robustness, there is no problematic circularity here: robustness appears at a higher level when different strands of evidence are aggregated. A related point worth mentioning is that the different ways of measuring or detecting should be at least minimally reliable – as Calcott (2011) points out, if the independent means have a probability of less than 0.5 of being correct, then adding more independent means does not increase the justification (or in Calcott’s terminology, likelihood of truth).

erroneous becomes increasingly small, but only if the ways are independent from each other. However, although the idea is intuitively plausible, the exact nature of this independence has turned out to be difficult to spell out, and remains a matter of debate.

There are various different approaches to the independence of evidence in the literature (e.g., Fitelson 2001, Franklin & Howson 1984, Kosso 1989, Staley 2004), and although none of them may be fully satisfactory, they do give some general outlines of a solution. First of all, it is clear that full-blown statistical independence is not what is required: different ways of measuring or detecting a property (such as temperature) will be statistically correlated, even when they are in other important respects independent (e.g., two thermometers based on different physical principles, such as a mercury thermometer and a radiation thermometer). A more promising notion is “less than perfect correlation” as proposed by Franklin and Howson (1984): if ways X and Y are less than perfectly correlated, then the hypothesis one wants to confirm receives more support from a mixture of X and Y than from repeated trials of X (or Y) alone (Franklin & Howson 1984). However, it is doubtful whether the nature of the required independence can be fully captured in formal terms. The idea is rather that the different ways of access should partly rely on different theoretical assumptions, different physical processes, or different experimental setups. What seems to be necessary (though not sufficient) is that any *problematic* or *unconfirmed* assumptions should not be shared by the different ways of access (cf. Stegenga 2012). All in all, this calls for a more detailed (article-length) account, which is beyond the scope of this paper – for more discussion on robustness and independence of evidence, see Franklin & Howson (1984), Nederbragt (2012) and Stegenga (2012).

Stegenga (2009) and Hudson (2014) have pointed out various other problems related to robustness as multiple accessibility. Stegenga (2009) raises issues such as dealing with discordant or conflicting evidence and evaluating the relative value or contribution of different strands (or types) of evidence. A further related issue is the integration or concordance of evidence of different types: the various modes of access

mentioned in the definition of robustness include scientific practices of quite different kinds (such as measuring and deriving), and it is not clear by what criteria we could ascertain that they provide access to the *same* phenomenon or entity.<sup>12</sup> Hudson (2014) is very critical of robustness, and goes through several cases of apparent robustness reasoning, arguing that what is epistemically important in these cases is not robustness but identifying a process that is sufficiently reliable (see also Soler 2014 for a detailed critical review of Hudson's book). However, rather than knock-down arguments, I take these to be issues that need to be clarified in future research, and constraints on a full-blown account of robustness, as also Stegenga (2012) is inclined to do.

#### 4. Robustness, justification, and science

In the beginning of section 3, I proposed that instead of seeing robustness as a criterion for what is real, we should see it as a source of justification for our ontological commitments. Now that we have discussed robustness in detail, we can return to this idea and spell it out more clearly. I propose to define the relation between robustness and ontological commitments as follows: *Robustness confers justification for believing that X is real, and the degree of this justification corresponds to the degree that X is robust.*

This thesis captures the idea that robustness is a source of justification for ontological commitments, and that the justification is a matter of degree. Importantly, it makes robustness a sufficient but not necessary condition: there may be other sources of justification. The fact that X is not robust does not imply that X is not real, or that we are not justified in believing X to be real. For example, phenomenal properties (the what-it's-like aspects of mental states) are not robust, but most philosophers of mind consider them to be real properties due to the (arguably) direct and immediate access

---

<sup>12</sup> I thank an anonymous referee for bringing this problem to my attention. The issue is also discussed in Hudson (2014) and Stegenga (2012).

that we have to them. It may also be that in some other special cases there is only one way of measuring or detecting some phenomenon, property or entity, but this way has been judged to be extremely reliable and error-free. In such cases we may be justified in believing that the phenomenon, property, or entity is real even though it is not robust.

In order to show that the account of robustness I am defending is supported by and reflects scientific practice, I will consider two examples from history or science, one drawn from the neuroscience of vision and the other from climate research.<sup>13</sup> The first example concerns “luminance units” in the retina. In the late 1960s, the neuroscientists Horace Barlow and William Levick were studying the responses of ganglion cells in the cat retina. Retinal ganglion cells form the last layer of processing in the retina, and are responsible for transmitting visual information to the brain. Barlow and Levick (1969) were measuring the electrophysiological responses of single ganglion cells, in order to find out how the base firing rate (“maintained discharge”) changes at different levels of light adaptation. To their surprise, they discovered units that were behaving very differently from other ganglion cells. The response of these cells was slow, relatively straightforwardly related to light intensity, and increased monotonically with increasing light intensity. These units were extremely rare – only three were positively identified out of a sample of hundreds of units. Barlow and Levick (1969) dubbed them “luminance units” and presented the tentative hypothesis that they convey information about luminance conditions.

At that time, this was all that was known about these luminance units. They had not been observed in other experimental setups. They had not been derived from any existing models, and there was no theoretical framework or model where these units could be incorporated into. Even the single way of measuring them was not very

---

<sup>13</sup> The classic example of robustness is Perrin’s argument for the existences of atoms based on multiple independent determinations of Avogadro’s number. I will not go through it here, since it has been discussed by several other authors (e.g., Cartwright 1984, Salmon 1984, Hudson 2014).



reliable due to the low number of units detected. Consequently, they were *not* robust (or robust to a very low degree), and researchers in the scientific community were not justified in believing that they are real. Indeed, Barlow and Levick (1969, 714) asserted that it is “obviously risky” to conclude that these units form a distinct category.

However, in recent years, a completely different line of research has led back to the luminance units (see Do & Yau 2010 for an extensive review). A type of rare and morphologically distinct ganglion cells has been discovered in the mammalian retina. These “intrinsically photosensitive retinal ganglion cells” (ipRGCs) contain a new type of photopigment, melanopsin. They respond to light stimuli even when all signals from the normal photoreceptors (rods and cones) are blocked – hence the title “intrinsically photosensitive”. Interestingly, ipRGCs have many of the same properties as luminance units: they are rare (comprising only about 1-3 % of the retinal ganglion cell population), they respond relatively slowly to light stimuli, the response increases monotonically with increasing light intensity, and the firing patterns of both types of units are similar. Thus, it is very likely that the luminance units discovered by Barlow and Levick were ipRGCs.

In contrast to luminance units in the 1960s, ipRGCs in the 2010s are very robust. There is a broad range of independent techniques and experimental setups with which they can be studied: immunostaining for melanopsin, single unit recording, patch clamp recording, gene knockout studies, and so on (Do & Yau 2010). They play an explanatory role in models of light adaptation and circadian photoentrainment (i.e., the synchronization of the “biological clock” that keeps track of the day-night cycle): in these models, ipRGCs convey information about the level of illumination (Kumar Nayak, Yegla, & Panda 2007). The projections of ipRGCs to the brain, most importantly to the lateral geniculate nucleus of the thalamus, have been extensively studied. All in all, ipRGCs are extremely robust, and this confers a very high degree of justification to the scientific community for believing that they are real. This in turn

allows for more sophisticated explanations and theories, since scientists have a more detailed and well-justified ontology of the retina at their disposal.

This example nicely illustrates the epistemic and relative nature of robustness: mammalian retinas contained ipRGCs in the 1960s just as they do now, but in the 1960s ipRGCs (or luminance units, as they were called then) were not robust, while now they are to a high degree. Thus, as I have argued in section 3, conceptions of what is robust, and the related degree of justification, change over time. Similar stories abound in the history of science: consider for example the photoelectric effect, which was first discovered by Hertz in the late 19<sup>th</sup> century, and was at first robust to a low degree only. However, it was quickly confirmed by further independent experiments, and finally explained and fully incorporated into the state-of-the-art physical theories by Einstein in 1905. For further case studies of robustness, see for example Nederbragt (2003) and Trizio (2012).

In the luminance unit case, considerations of robustness remain largely implicit – for example, Do & Yau (2010, 1547-1551) start their review by going through the different ways in which ipRGCs have been detected and measured, but do not explicitly draw the conclusion that this justifies us holding them to be real, probably because in the light of the evidence this conclusion is nearly trivial. However, in more controversial cases considerations of robustness are also explicitly appealed to.

Let us take a brief look at one such case: climate change (see Trizio 2012 for another example from cell biology). Since the reality of global warming is still doubted by some non-scientists, the reports of the Intergovernmental Panel for Climate Change (IPCC) often refer to robustness to justify the conclusions. Interestingly, the fourth IPCC synthesis report even uses the term itself, defining it in a way that comes close to robustness as multiple accessibility: “a robust finding for climate change is defined as one that holds under a variety of approaches, methods, models and assumptions, and is expected to be relatively unaffected by uncertainties” (IPCC 2007, 72).<sup>14</sup> This

---

<sup>14</sup> This definition is not explained in more detail, and thus could also be understood as something closer to DR. However, the example that follows (warming of the

is followed by a list of the key robust findings, such as “[w]arming of the climate system is unequivocal, as is now evident from observations of increases in global average air and ocean temperatures, widespread melting of snow and ice and rising global average sea level.” Thus, climate scientists appeal to the variety of different ways of measuring, deriving, or detecting the warming of the climate system in order to justify the belief that the phenomenon is real.

These examples show that robustness plays an important role in justifying ontological commitments in science (either implicitly or explicitly), in accordance with the account I have defended in this paper. In the next section, I turn to the relevance of robustness for general issues in philosophy of science.

## 5. Robustness and scientific realism

The approach defended in this paper – taking robustness as a source of justification for ontological commitments – has the potential to shed light on various topics in philosophy of science. Most importantly, as I will show in this section, the robustness approach opens up new perspectives to issues related to scientific realism. The aim here is not to give a full defense of any particular position, but to illustrate some of the ways in which robustness can be put to use.

The idea defended above is that we are justified in believing that things studied by science are real insofar as they are robust. Understood in this way, robustness fulfils an important and neglected role in contemporary philosophy. As I pointed out in the introduction, philosophers have traditionally focused on searching for a metaphysical criterion or definition for what is real, and have ended up defending principles such

---

climate system) clearly exhibits robustness as MA, as it refers to “observations” and not just models. In any case, the main purpose of this brief example is to demonstrate the importance of robustness for scientific debates. It is not intended to further clarify the notion, or to provide an analysis of the exact kind of robustness reasoning applied by the IPCC (which might be an interesting project in its own right).

as “to be real is to have causal powers” (e.g., Kim 1993, 202)) or “all real things are composed of fundamental physical particles” (e.g., Pettit 1993).<sup>15</sup> These criteria are not relativized to a specific scientific community or point in time, and answer questions pertaining to the objective, final, and fully mind-independent make-up of the world. However, there is a different but related philosophical question that has received much less attention, and to which the traditional accounts are largely irrelevant: What are we justified in holding to be real now, as limited beings, based on the current state of science? Robustness provides one answer to this question.

It is also possible to take the robustness approach a step further, arguing that the kind of justification based on robustness is the best that we can have, and that further inferences to what is *really* real are problematic and plagued with problems such as the pessimistic induction. In other words, any attempts to define what is real and non-real in a metaphysically deeper way are doomed to fail: there is a mind-independent reality, but we can never be certain of having epistemic access to its fundamental nature. This (somewhat Kantian) basic idea has a long tradition, and is defended in different forms by Ladyman & Ross (2007), Quine (1981) and Wimsatt (2007) (see also Boon (2012), who defends a broadly Kantian approach to science and reality based on various robustness notions).

A closely related view is Arthur Fine’s (1984) “Natural Ontological Attitude” (NOA). Fine argued that empirically justified everyday “homely truths” are on a par with scientific truths, and that we should accept the truth of the latter just as we accept the truth of the former – trying to provide a philosophical interpretation of the truth of scientific statements or of the correspondence between scientific theories and reality is misguided and unnecessary. Fine named the position that simply accepts the truth of the results of science and everyday truths “the core position”, and argued that we

---

<sup>15</sup> I do not intend to dismiss such criteria, but rather to point out that they have a different role than robustness. In fact, these criteria are entirely compatible with the robustness approach. For example, one could argue that the ultimate metaphysical criterion for what is real is the causal criterion, but the source of justification for science-based ontological commitments is robustness.

should simply accept it as such without any of the additional claims that realists or antirealists make. This in turn commits us to the existence of the entities and properties that scientific theories refer to. The robustness approach can be used to support this view, as it shows that we are justified in believing that the objects studied by science are real, regardless of any additional metaphysical issues concerning the nature of mind-independent reality and its relation to scientific theories.

However, supporters of stronger forms of realism may also benefit from a closer analysis of robustness. It can be used to formulate the following arguments in favor of scientific realism. First of all, many entities and phenomena, for example electrons, have remained robust throughout much of the history of modern physics, which could be taken as evidence for their mind-independent reality (cf. the “entity realism” of Hacking (1982) and Cartwright (1983)). Furthermore, it could be argued that the robustness of phenomena studied by science tends to increase over time. It is plausible that scientific and technological progress, generally speaking, results in more and more independent ways of accessing entities, properties, and phenomena. For example, there are many more different ways of detecting, measuring, or deriving electrons now than one hundred years ago, and the same holds for most things studied in contemporary physics. If this metascientific hypothesis turned out to generally hold, it would mean that our confidence or justification for believing in the reality of things studied by science increases over time.

One apparent counterexample to this is cases like phlogiston, which may have been robust to some degree at some point in time, but later turned out not to exist at all (cf. Enc 1976). Prima facie, such cases suggest that robustness does not in general increase over time. I have two responses to this concern. First, it could be argued that phlogiston was never robust, since all the different ways of deriving, producing, or detecting it were either based on (shared) unsupported theoretical assumptions, or where flawed in other ways. That is, although phlogiston seemed to be accessible in multiple independent ways, this was not actually the case (this would be comparable to the classic case of pseudo-robustness, bacterial mesosome – see Culp (1994) for

more). Secondly, even if we grant that phlogiston was at some point in time robust to some low degree, it was clearly replaced by a far more robust entity, namely oxygen. Thus, also in cases like this there is a general increase over time of the robustness of the phenomena studied by science. The scientific ontology may be refined and revised over time, but the overall confidence in it increases.

Apart from these issues, a further implication of the robustness approach is that it supports our belief in the reality of higher-level things. As the examples discussed in the previous section suggest, there is no reason to doubt that many or most of the entities, properties and phenomena studied by the special sciences are robust, which implies that we are justified in believing that they are real. However, this does not lead to any ontologically radical pluralism or emergentism. The robustness approach shows that we can have direct justification for holding higher-level things to be real, but this is entirely compatible with most forms of (non-eliminative) reductionism. The reality of higher-level things leaves open the possibility that they can be mechanistically explained (see Bechtel 2008, Eronen 2012 and Wimsatt 2007), or even that they may be in some sense ontologically reducible (i.e. identical) to lower-level things (e.g. Esfeld & Sachse 2007).

Interestingly, the robustness approach may lead to the striking conclusion that we are, in some cases, *more* justified in believing that higher-level things are real than lower-level things. For example, neurons and biological organisms are extremely robust, while some fundamental physical particles such as the Higgs boson are (as of yet) robust to a relatively low degree. This of course does not mean that fundamental physical particles are less real than neurons, but rather that our justification for holding them to be real is relatively lower: because of the difficulties related to measuring and examining elementary particles, it is more likely that the Higgs boson turns out not to exist than that neurons turn out not to exist.

A staunch antirealist (along the lines of van Fraassen 1980) would probably not be satisfied with the above, and could simply deny that robustness (or anything else for that matter) provides justification for holding things studied by science to be real.

However, my aim here is not to give a full-blown defense of scientific realism, but to show how robustness provides new ways of approaching the issue and new possible arguments. I also believe that antirealists could happily accept the importance of robustness and many of the associated claims (see Boon 2012). The definition of robustness itself does not say anything about reality, and the part about justification could probably be reformulated in an antirealist framework (e.g., in terms of empirical adequacy and acceptance instead of ontological commitments).

## 6. Conclusions

Robustness is often presented as giving guidelines for distinguishing the true or real from mere appearances or artifacts. Most of recent discussions of robustness have focused on the kind of derivational robustness analysis introduced by Levins. The extent to which it justifies inferences to what is real is far from clear and a matter of ongoing debate. Instead, I have focused in this paper on the idea of robustness as multiple accessibility, drawn from the work of Wimsatt. I have developed and analyzed this idea, and argued that it can provide justification for ontological commitments. The idea is that we are justified in believing that things studied by science are real insofar as they are robust. As I have shown, this idea can be useful to philosophers in science in many ways, especially in debates related to scientific realism. Most importantly, as opposed to more metaphysical criteria, it shows what we are justified in holding to be real now, as limited beings, based on the current state of science.

### **Compliance with Ethical Standards**

The research resulting in this paper was funded by the Research Foundation Flanders – FWO (Postdoctoral Fellowship).

## Acknowledgments

I would like to thank the (five) anonymous referees of this journal, whose insightful and detailed comments were extremely useful in improving the manuscript. I am also very grateful to the following individuals for their helpful comments on earlier drafts: Hugh Desmond, James DiFrisco, Harmen Ghijsen, Chris Kelp, Jaakko Kuorikoski, Jani Raerinne, Paul Teller, and Raphael van Riel, as well as audiences at Ruhr University Bochum, University of Groningen, and KU Leuven. I especially thank Jan Heylen and Laura Bringmann for their very constructive and helpful feedback on several versions of the paper.

## References

- Barlow, H. B., & Levick, W. R. (1969). Changes in the maintained discharge with adaptation level in the cat retina. *Journal of Physiology*, 202, 699-718.
- Bird, A. (1998). *Philosophy of Science*. London: Routledge.
- Boon, M. (2012). Understanding Scientific Practices: The Role of Robustness Notions. In Soler et al. (2012) (pp. 289-315).
- Boyd, R. (1991). Realism, Anti-Foundationalism and the Enthusiasm for Natural Kinds. *Philosophical Studies*, 61, 127-148.
- Calcott, B. (2011). Wimsatt and the robustness family: Review of Wimsatt's Re-engineering Philosophy for Limited Beings. *Biology & Philosophy*, 26, 281-293.
- Cartwright, N. (1983). *How the Laws of Physics Lie*. Oxford: Clarendon Press.
- Cartwright, N. (2007). *Hunting Causes and Using Them. Approaches in Philosophy and Economics*. Cambridge: Cambridge University Press.
- Chang, H. (2004). *Inventing Temperature. Measurement and Scientific Progress*. Oxford: Oxford University Press.



- Crupi, V., Tentori, K. & Gonzalez, M. (2007). On Bayesian Measures of Evidential Support: Theoretical and Empirical Issues. *Philosophy of Science*, 74, 229-252.
- Culp, S. (1994). Defending Robustness: The Bacterial Mesosome as a Test Case. In D. Hull, M. Forbes & R. M. Burian (Eds.) *Proceedings of the Biennial Meeting of the Philosophy of Science Association 1994, Vol. 1* (pp. 46-57). East Lansing: The Philosophy of Science Association.
- Do, M. T. H., & Yau, K.-W. (2010). Intrinsically Photosensitive Retinal Ganglion Cells. *Physiological Reviews*, 90, 1547-1581.
- Enç, B. (1976). Reference of Theoretical Terms. *Noûs*, 10, 261-282.
- Eronen, M. I. (2012). Pluralistic Physicalism and the Causal Exclusion Argument. *European Journal for Philosophy of Science*, 2, 219-232.
- Esfeld, M. & Sachse, C. (2007). Theory Reduction by Means of Functional Sub-Types. *International Studies in the Philosophy of Science*, 21, 1-17.
- Fine, A. (1984). The Natural Ontological Attitude. In J. Leplin (Ed.), *Scientific Realism* (pp. 83-107). Berkeley: University of California Press.
- Fitelson, B. (2001). A Bayesian Account of Independent Evidence with Applications. *Philosophy of Science* 68, S123-S140.
- Forber, P. (2010). Confirmation and explaining how possible. *Studies in History and Philosophy of Biological and Biomedical Sciences*, 41, 32-40.
- Franklin, A. & Howson, C. (1984). Why Do Scientists Prefer to Vary Their Experiments? *Studies in the History and Philosophy of Science*, 15, 51-62.
- Hacking, I. (1982). Experimentation and Scientific Realism. *Philosophical Topics*, 13, 154-172.
- Hacking, I. (1983). *Representing and intervening. Introductory topics in the philosophy of natural science*. New York: Cambridge University Press.
- Hudson, R. (2014). *Seeing Things: The Philosophy of Reliable Observation*. Oxford: Oxford University Press.
- IPCC (2007). Core Writing Team, Pachauri, R.K and Reisinger, A. (Eds.), *Climate Change 2007: Synthesis Report. Contribution of Working Groups I, II and III to the Fourth Assessment Report of the Intergovernmental Panel on Climate Change*. Geneva, Switzerland: IPCC.

- Kim, J. (1993). The non-reductivist's troubles with mental causation. In J. Heil & A. Mele (Eds.), *Mental Causation* (pp. 189-210). Oxford: Clarendon Press.
- Kumar Nayak, S., Jegla, T., & Panda, S. (2007). Role of a novel photopigment, melanopsin, in behavioral adaptation to light. *Cellular and Molecular Life Sciences*, 64, 144-154.
- Kuorikoski J., Lehtinen A., & Marchionni C. (2010). Economic modelling as robustness analysis. *British Journal for the Philosophy of Science*, 61, 541–567.
- Kuorikoski, J., Lehtinen, A., & Marchionni, C. (2012). Robustness analysis disclaimer: please read the manual before use! *Biology & Philosophy*, 27, 891-902.
- Ladyman, J., & Ross, D. (2007). *Every Thing Must Go: Metaphysics Naturalised*. Oxford: Oxford University Press.
- Laudan, L. (1981). A Confutation of Convergent Realism. *Philosophy of Science*, 48, 19-48.
- Levins, R. (1966). The strategy of model building in population biology. *American Scientist*, 54, 421-431.
- Matthewson, J., & Weisberg, M. (2009). The structure of tradeoffs in model building. *Synthese*, 170, 169-190.
- Nederbragt, H. (2003). Strategies to improve the reliability of a theory: the experiment of bacterial invasion into cultured epithelial cells. *Studies in History and Philosophy of Biological and Biomedical Sciences*, 34, 593-614.
- Nederbragt, H. (2012). Multiple Derivability and the Reliability and Stabilization of Theories. In Soler et al. (Eds.) (pp. 121-145).
- Odenbaugh, J., & Alexandrova, A. (2011). Buyer beware: robustness analyses in economics and biology. *Biology & Philosophy*, 26, 757-771.
- Orzack, S. H., & Sober, E. (1993). A critical assessment of Levins's The Strategy of Model Building in Population Biology (1966). *The Quarterly Review of Biology*, 68, 533–546.
- Pettit, P. (1993). A definition of physicalism. *Analysis*, 53, 213-223.
- Raerinne, J. (2013). Robustness and sensitivity of biological models. *Philosophical Studies*, 166, 285-303.
- Quine, W. v. O. (1981). *Theories and Things*. Cambridge, MA: Harvard University Press.

- Salmon, W. C. (1984). *Scientific Explanation and the Causal Structure of the World*. Princeton, NJ: Princeton University Press.
- Soler, L., Trizio, E., Nickles, T., & Wimsatt, W. C. (Eds.). (2012). *Characterizing the Robustness of Science: After the Practice Turn in the Philosophy of Science*. Dordrecht: Springer.
- Soler, L. (2014). Against Robustness? Strategies to Support the Reliability of Scientific Results. *International Studies in the Philosophy of Science*, 28, 203-215.
- Staley, K. W. (2004). Robust Evidence and Secure Evidence Claims. *Philosophy of Science*, 71, 467-488.
- Stegenga, J. (2009). Robustness, Discordance, and Relevance. *Philosophy of Science*, 76, 650-661.
- Stegenga, J. (2012). *Rerum Concordia Discors: Robustness and Discordant Multimodal Evidence*. In Soler et al. (2012) (pp. 207-226). Trizio, E. (2012). Achieving Robustness to Confirm Controversial Hypotheses: A Case Study in Cell Biology. In Soler et al. (Eds.) (pp. 105-120).
- Trout, J. D. (1998). *Measuring the Intentional World*. Oxford: Oxford University Press.
- Van Fraassen, B. (1980). *The Scientific Image*. Oxford: Oxford University Press.
- Weisberg, M. (2006). Robustness analysis. *Philosophy of Science*, 73, 730-742.
- Weisberg, M., & Reisman, K. (2008). The robust Volterra principle. *Philosophy of Science*, 75, 106-131.
- Wimsatt, W. C. (1981). Robustness, reliability, and overdetermination. in M. Brewer, & B. Collins (Eds.) (1981), *Scientific Inquiry and the Social Sciences* (pp. 124-163). San Francisco: Jossey-Bass. Reprinted in W. C. Wimsatt (2007) *Re-Engineering Philosophy for Limited Beings. Piecewise Approximations to Reality* (pp. 43-74). Cambridge, MA: Harvard University Press.
- Wimsatt, W. C. (1994). The ontology of complex systems: levels of organization, perspectives, and causal thicketts. *Canadian Journal of Philosophy* S20, 207-274. Reprinted in W. C. Wimsatt (2007) *Re-Engineering Philosophy for Limited Beings. Piecewise Approximations to Reality* (pp. 193-240). Cambridge, MA: Harvard University.

Woodward, J. (2003). *Making Things Happen. A Theory of Causal Explanation*. Oxford: Oxford University Press.

Woodward, J. (2006). Some varieties of robustness. *Journal of Economic Methodology*, 13, 219–240.