Chapter 7
Scientific Images and Robustness

Catherine Allamel-Raffin and Jean-Luc Gangloff

As Léna Soler[1] has emphasized:

> (. . .) the term 'robustness' (. . .) is, today, very often employed within philosophy of science in an intuitive, nontechnical and flexible sense that, globally, acts as a synonym of 'reliable', 'stable', 'effective', 'well established', 'credible', 'trustworthy', or even 'true'.

But in parallel, William C. Wimsatt has developed a specific sense (Wimsatt 1981, 2007), which, while preserving the common association with the ideas of reliability and effectiveness, is more precise and more technical, and refers to the idea of the invariance of a result under multiple independent derivations. In this paper, we argue that "robustness analysis" (in Wimsatt's sense) is nothing less than the guiding principle of the argumentative structure of many papers published in natural sciences. We base our analysis on the methodology of ethnographic studies. Our aim is to take into account the actual practices which occur in laboratories (Allamel-Raffin 2004, 2005).[2] Our approach is mainly descriptive, although it does not exclude a normative perspective; for we conceive norms to be elaborated in the research process. In other words, we believe that norms are historically set up. Besides, we think that problems raised by philosophers are also faced by scientists. For example, "the experimenter's regress" or "the theoretical underdetermination by the data", as Kitcher (2001) observed are not only issues identified by philosophers but also by scientists themselves. Our argument is based on the examination of a 2001 astrophysical paper: "The Milky Way in molecular clouds: a new complete CO survey".[3] One of our purposes is to link the "robustness analysis" with the use of images in scientific papers. We shall see how images can never be reduced to mere illustrations but are an important component of the argumentation, and thereby

---

[1] See Chapter 1, p. 3.

[2] Our study relies on an ethnographic investigation conducted in an astrophysics lab: the Harvard-Smithsonian Center for Astrophysics (CfA), Cambridge, USA.

[3] Dame et al. (2001).

C. Allamel-Raffin (✉)
IRIST, University of Strasbourg, Strasbourg, France
e-mail: allamelraffin@unistra.fr

L. Soler et al. (eds.), *Characterizing the Robustness of Science*, Boston Studies in the Philosophy of Science 292, DOI 10.1007/978-94-007-2759-5_7,

of the robustness of the results. First, we shall focus on the various discourses on robustness in philosophy of science in order to show that Wimsatt's "robustness analysis" concept is better able to capture laboratory practices but also the structure of a scientific paper (Sections 7.1 and 7.2). In a second part, we study how generally a scientific paper is not a demonstration, strictly speaking, but rather a sequence of arguments (Section 7.3). For that purpose, we examine the argumentation of the aforementioned astrophysical paper and especially the role played by images in the argumentation (Sections 7.4, 7.5, and 7.6). Finally, insofar as the notions of independence and invariance are crucial to the concept of "robustness analysis", we shall comment on them relying on our case study on the Milky Way (Section 7.7).

7.1 Robustness Analysis of Philosophers

The concept of "robustness" has been explicitly thematized in philosophy of science for three decades. As with any concept, it's a "working" concept, and we can use it in different contexts. And as with most empirical concepts, this one has an "open texture", as argued by Waismann (1945). Empirical concepts are not defined as a set of characteristics established once and for all (as in the formal sciences). They change over time, and one can imagine that their new characteristics are influenced by new practices or new conceptual frames. This is probably the reason why "robustness" has generated different definitions. Several authors have proposed different versions of the concept of "robustness" and illustrated them by different examples. For example, in a case study of DNA sequencing, Culp (1995) shows how two completely different methods led to comparable results. Nederbragt (2003) presents a case study of the invasion of cells by microorganisms. He also shows that scientists have used two or three different methods or instruments in order to confirm their theory. In a historical study about the reliability of thermometers, Chang (2001, p. 283) explains that the use of several thermometers – based on different principles: air, carbonic acid, hydrogen – led the physicist Regnault to close the debate about temperature. Chang insists on showing that in this case there was no appeal to any theory to close the discussion. Ian Hacking, for his part (1981, pp. 144–145; 1983, pp. 324–332) takes the example of the dense bodies in red blood platelets. He shows that these entities can be detected by two different microscopes relying on different properties of light, namely, the transmission electron microscope and the fluorescent microscope. Hacking[4] terms this "the argument from coincidence".

---

[4] He also develops another concept: the "robust fit". In a sense, this concept is close to robustness analysis. What is a "robust fit"? We will mainly rely on the definition proposed by Hacking (1992). According to this view, the "robust fit" is the adjustment, in the laboratory, between three fundamental categories (ideas, things and marks) that enable scientists to obtain the reliability and the repeatability of the results. The concept of "robust fit" is slightly different from the robustness analysis, but what is in common, is the procedure consisting in *crossing several elements* (theories, instruments, know-how, etc.) to obtain robust scientific results.

Taking into account all these analyses, Nederbragt affirms that several names were given to the same procedures consisting in 'using different methods to confirm an hypothesis'. These denominations can be: "robustness analysis" (Wimsatt 1981), "triangulation" (Culp 1995; Wimsatt 1981), and "independence of the routes" (Hudson 1999). In our eyes, all these definitions have something in common: they fit Wimsatt's definition, which is:

> The family of criteria and procedures which I seek to describe in their various uses might be called *robustness analysis*. They involve the following procedures: 1/ To analyse a *variety* of *independent* derivation, identification, or measurements processes. 2/ To look and analyse things which 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 analyse and explain any relevant *failures* of *invariance*. I will call things which are invariant under this analysis "robust". (Wimsatt 1981, p. 126)

For us, this definition is by far the more precise and technical. It's the reason why we will rely our analysis on it.

7.2 Robustness Analysis in Scientific Laboratories

What was extremely striking during our ethnographic studies is how the robustness analysis is entrenched in the day-to-day activities of an astrophysical laboratory such as the Center for Astrophysics in Harvard. Robustness analysis is a fundamental methodological principle because the astrophysicists' enquiry has to deal with two dimensions of observability: directness and amount of interpretation (Kosso 1989). According to Peter Kosso, directness can be understood as a dimension of observability: there are more or fewer interactions from the source X observed conveying an information to the final human receptor. Directness, according to Kosso, is a measure of the physical closeness between the source and the final receptor. In contemporary astrophysics, the chain of interactions included in an observation is often long and complex and the observation consequently hugely indirect. Another dimension of observability is the amount of interpretation: how many distinct physical laws are needed to get from the source X to the final receptor. The amount of interpretation is a measure of epistemic closeness between this source and the final receptor. The number of such laws, in astrophysical observations, is often very high. In absolute terms, for one observation, the scientists should be able to make explicit what kind of physical interactions occurred and which physical laws are involved in the observation. They should also be able to make explicit how they identified and quantified the noise. What makes things a lot more difficult is the possible existence of artefacts (the sources of certain of these artefacts being unknown) and the existence of tacit knowledge due to the fact that this information is produced by human beings. To reduce these difficulties, one can think that the solution must be found in instrumentation, in the sense that instruments perform a lot of tasks automatically. But that would be too simple… The old ninenteenth century ideal of "mechanical objectivity" (Daston and Galison 1992, 2007) and consequently the hope that scientists' subjectivity could be eliminated, has vanished. We know that we can't avoid the presence of human subjectivity in the use of instruments. We know that the use of each instrument includes subjective choices and also skills, the canonical example being in this case photographical techniques. Idiosyncratic characteristics are always present and cannot be definitively eliminated. The only way to get out of this impasse is to recognize that objectivity comes in degrees (Culp 1995) and has to be conceived more as a *continuum* (Putnam 2003). For those who agree with this view, the goal is then to reduce idiosyncratic characteristics as much as possible.

In the end, we can recognize especially if we stay a long time in a lab,[5] that we have an interpretative flexibility of the data as Harry Collins (1981) describes in his papers, but:

a) we can notice that the scientists are perfectly aware of this interpretative flexibility of the data;
b) Therefore, the scientists try to reduce this interpretative flexibility. The robustness analysis seems for them a *practical necessity.* They aim above all to produce 'robust' results.

7.3 Argumentation in Scientific Papers

It follows from these ideas that numerous scientific papers do not "demonstrate", but propose an argumentation. The astrophysicists get with their data or images only what we can call "pieces of evidence" or "elements of proof". They do not start with true premises to get in the end, by following the rules of formal logic, absolutely true conclusions. Thereby, they are forced to propose in their papers the more convincing way to expose their pieces of evidence or elements of proof. What is argumentation? Briefly, it includes the following features[6]: any argument contains a claim, an assertion put forward publicly for general acceptance. This claim is supported by grounds or statements specifying particular facts about a situation. These facts are already accepted as true and can therefore be relied on to clarify and make good the previous claim (establish its truth or correctness). Sometimes, we need warrants, statements indicating how the grounds or facts on which we agree are connected to the claim. These connecting statements draw attention to the previously agreed general ways of arguing applied to the particular case and so are implicitly relied on

---

[5] This conclusion is drawn from our ethnographic studies. We spent three months in the Harvard-Smithsonian Center for Astrophysics (USA), six months in a nanoscience laboratory (France) and three months in a pharmacology laboratory (France). In each case, we observed numerous times this phenomenon of interpretative flexibility, especially when the scientists work on scientific images. For example: see Allamel-Raffin (2005).

[6] Our main references concerning the question of argumentation are the following books: Toulmin (1958) and Janik et al. (1984). We are perfectly aware that a lot of literature exists on the argumentation topic, but we have chosen to use the concepts of Toulmin & *al*., because they are useful enough to help us to understand the general argumentation of a scientific paper.

as those whose trustworthiness is well established. In scientific papers, the critics can focus on the grounds, the data itself (arguing that it is false). But they can also ask questions about the use of a given instrument or a given technique employed to obtain this data. Sometimes, the warrant has to be justified, and that is the function of the backing, the generalization making explicit the body of experience relied on to establish the trustworthiness of the ways of arguing applied in any particular case. With this definition, we'll examine in our two next sections the astrophysical paper's argumentation.

7.4 A Case Study[7]: A New Complete CO Survey in the Milky Way

The paper's title is: "The Milky Way in molecular clouds: a new complete CO survey". The paper is a perfect illustration of the work done in a context of normal science activity. It's a radio astronomy[8] paper, one of the most cited papers in radio astronomy since it was published in 2001. The article includes two big maps (1.5 m long). These maps constitute the main result of the paper and represent the CO distribution in our galaxy (Fig. 7.1).

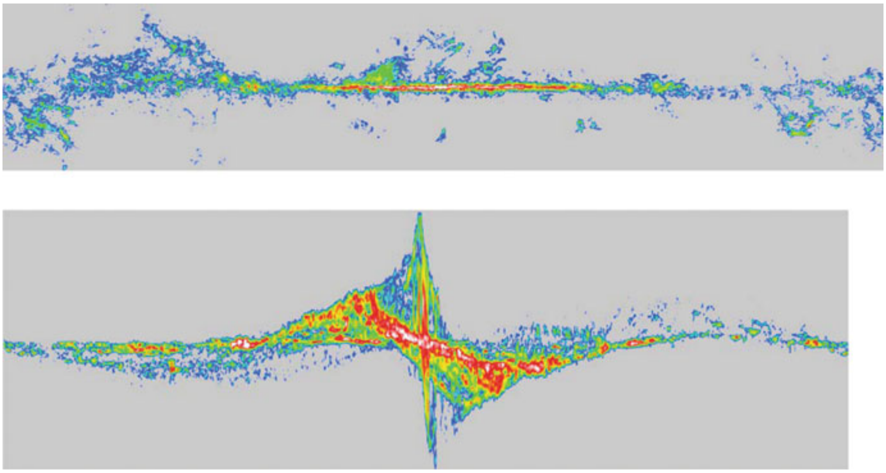

Fig. 7.1 Two big radio maps of the Milky Way (1a and 1b)
Source: Dame et al. (2001, CfA)

---

[7] Our analysis relies not only on the paper itself. It also includes some ethnographic material collected during our stay in the Center for Astrophysics at Harvard: interviews recorded with one of the paper's authors (Thomas Dame) and observation reports of the day-to-day activities in the lab.

[8] Radio astronomy is a subfield of astronomy that studies celestial objects in the radio frequency portion of the electromagnetic spectrum, that is to say the wavelength between 0.3 and 2500 mm.

Why is it interesting to study the presence of CO in the Milky Way? In fact, CO detection is a mean to study dust clouds. These dust clouds are very interesting for many reasons. One of these reasons is that they are the birthplace of new stars. Dust clouds are made of molecular gas and atomic gas. There are many molecular gases, H_2 being the most important. But H_2 is extremely difficult to detect from the earth. CO is also a molecular gas, and many studies show that it's mixed with H_2 in dust clouds. So CO is a tracer of H_2 and of the dust clouds in the interstellar medium. CO is relatively easy to detect (its wavelength is 2.6 mm). The purpose of the two maps in the paper is to be useful for many other studies such as the aforementioned studies of the birth of stars, studies of the source of the cosmic rays, studies on the structure of our galaxy, etc.

Two associated claims are defended in the paper: the CO maps of the Milky Way:

(1) are constituted of reliable data
(2) are complete (there is no lack of data concerning the Milky Way).

The scientists want to argue that they have good quality data (no noise, no artefact). In other words, they argue that the data in the maps corresponds to reliable data. In the case, the data is about CO clouds. Furthermore, they argue that their two maps are complete (a map with some missing pieces would be useless). Our first preliminary analysis will focus on the arguments presented to support those two claims. What are the arguments exposed in the subsections of the paper in order to defend what we call the 'reliability claim'?

In Sections 7.4.1 and 7.4.2, we shall briefly analyze the structure of the paper and show that it fits very well with the robustness analysis of W. C. Wimsatt.

7.4.1 The Reliability Claim

7.4.1.1 First Argument: Analysis of the Instrumentation and of the Data Processing in Order to Justify Their Reliability

The paper begins with a description of the two telescopes used, including a brief history, and a discussion of the data acquisition and reduction employed in the various surveys. The scientists detail carefully the calibration procedures. A central point is constituted by the synthesis of the data. The survey was constructed from 37 individual surveys. It was crucial for the astrophysicists to explain how they managed to synthesize all these surveys, and more specifically, how they managed to reduce the noise included in these different studies. The way they suppressed noise is especially explained here because they didn't use the usual procedure to reduce noise in radio astronomy. If we compare the selected paper with other papers in astrophysics, this section is more developed than the same section in a current astrophysics paper. Briefly, the telescopes are two small millimeter-wave telescopes (1.2 m telescopes), one at the CfA (Cambridge Massachussetts), the other at the Cerro Tololo InterAmerican Observatory in Chile.

About the data acquisition, a first survey was realized between 1979 and 1987. This survey was published in 1986. The data acquisition[9] went on until 2000, always using the same two telescopes. The results (the maps) of the first survey were integrated to the new survey. The final result here is the CO maps. Theses maps are in fact the combination of 37 different small surveys. How did the scientists get their data ? It was a meticulous activity lasting more than 20 years. One of the researchers told us what they actually did in order to obtain their data. We quote him:

> So, we look at one spot in the sky at a time in order to build up these images. Each observation, because the signal is very weak, typically takes 2 or 3 minutes. So, that is why it takes such a long time.

The result of these observations is data cubes, in other words files in FITS format. The data files tell you all you need to know: when the data was taken, what part of the sky, what range in velocity, what frequency, etc. If the data stays in that format, it would be unexploitable. This raw data has to be converted into images.

Once the data collected with the FITS format,[10] the scientists have to manage the noise. Because they wanted to put all 37 surveys into one map, the sensitivity should have been limited by the worst survey, and in particular by the noise of this worst survey. Generally, in radio astronomy, astrophysicists show a little bit of noise on the map because human eyes are good at picking out real things from noise. Sometimes, the most interesting features are almost in the noise. One way to deal with the noise and to reduce it is to smooth the maps. To smooth a map means to take, for each pixel, a weighted average of the pixels around. The same thing must be done for each pixel. As a consequence, the researcher gets a fuzzier image, but it is more sensitive to anything that is extended. On the other hand, he can lose some very strong sources. That is a problem because what the scientists want in our case is to obtain a map – including very weak and very strong CO radio sources. So they couldn't use the traditional method of reducing noise. In order to resolve this problem, the astrophysicists employed a technique called "moment analysis". They took the whole data cube, and they first smoothed it quite heavily. What they did then was to degrade the resolution. This gave them a greater sensitivity for anything that is extended. They then used this smoothed data cube to reduce the noise: in any place in the data cube where there is no emission or where they think there is no significant emission, they blanked it in the original data cube. Basically, they used the smoothed data cube to tell them the regions of the data cube where there might be real emission

---

[9] In this section of the article, a special part is dedicated to the calibration procedures. In order to be sure that their calibration was correct, the scientists chose two different methods of calibration. These calibration choices are justified by referring to other studies in radio astronomy. These calibration procedures have nothing special in regard of the calibration procedure used in other radio astronomy studies of the same kind.

[10] FITS stands for 'Flexible Image Transport System' and is the standard astronomical data format endorsed by both NASA and the IAU. FITS is much more than an image format (such as JPG or GIF) and is primarily designed to store scientific data sets consisting of multi-dimensional arrays (1-D spectra, 2-D images or 3-D data cubes) and 2-dimensional tables containing rows and columns of data. http://heasarc.gsfc.nasa.gov/docs/heasarc/fits.html.

or significant emission. And where there is no emission, they just blanked out every pixel in the original data cube. What they got was a moment cube, which had the original resolution as they hadn't degraded the resolution, but the noise was reduced. After that, they could integrate all the way through for all velocities without picking up much noise. Thus they obtained a map where the noise is reduced and where the sensitivity is non-uniform. All this required an enormous amount of work! After that the radio astrophysicists checked many of the weaker features of the integrated maps to assure that they corresponded to identifiable spectral lines in the corresponding spectra and not to base-line fluctuations or statistical noise.

After these calibration procedures and data processing, why did the scientists persist in their activities? At this point they had thoroughly detailed the conditions of the data collected and the way they processed it. Isn't that enough to convince the readers of the paper? The claim that the points on the maps correspond to reliable data is supported by the description they made of the instrumentation, data acquisition and reduction. But that's not enough.

The warrant isn't strong enough: the warrant, here, consists in asserting that the instrumentation and the data processing are reliable. But only the quality of the data could certify these points. So, we are in the experimenter's regress… (Collins 1992; Collins and Pinch 1993). The data is good if the instrumentation and the data processing are correct, but to certify that the instrumentation and the data processing were correctly done, you have to look at the data!

In fact, as Peter Winch says, the scientific investigator is involved in two sets of relations: first, with the phenomena he investigates and second with his fellow-scientists and the rules of his scientific community (Winch 2007, p. 84). The authors of this paper know perfectly well that even if they are confident in their result because the instrumentation and the data processing are correct, it will be not enough to convince their fellows. They know that because they have internalized that culture through training in their field. As Gingras and Godin (2002) notice, "this is why scientists can anticipate criticism" and in our case, they know perfectly that good use of instrumentation and correct data processing is far from being a sufficient argument to convince their peers. It is why the argumentation must go on: four arguments, all related to the robustness analysis, are presented.

7.4.1.2 Second Argument for the Reliability Claim

In the same region of the Milky Way, the astrophysicists compare their data with other data obtained with a much bigger radiotelescope, with a ten times better resolution. The name of this telescope is the FCRAO.[11] This study was carried out by

---

[11] FCRAO stands for Five College Radio Astronomy Observatory. It was founded in 1969 by the University of Massachusetts, together with Amherst College, Hampshire College, Mount Holyoke College and Smith College.

Heyer et al. in 1998. The data is concordant. Here we have an independent[12] survey, but the same type of instrumentation with better resolution.

7.4.1.3 Third Argument for the Reliability Claim

This argument is a comparison of the data obtained by the astrophysicists on well known celestial regions as Orion, Centaurus, etc. with data obtained on the same regions by other teams of scientists using either radiotelescopes or other telescopes like optical, X-rays, gamma rays telescopes. Here we have several independent surveys[13] and various instruments characterized by independent physical principles.

7.4.1.4 Fourth Argument for the Reliability Claim

Finally, they compare their data pertaining to the center of the galaxy with data obtained with an optical telescope. Indeed, the CO is one of the gases present in the dust clouds. Other studies showed that dust blocks the distant starlight emitted in the optical wavelength. If they took an image of the center of the Milky Way in optical wavelength, they should observe obscure areas corresponding to the dust clouds made partially of CO gas. In this case we have an independent survey made by various instruments characterized by independent physical principles.

7.4.2 The Completeness Claim

The scientists aim to present a ‘complete’ CO survey of our galaxy, that is to say, they ignore only an infinitesimal part of the CO existing in our galaxy. To support that claim, they appeal to other studies of the galaxy:

- the first is a radio survey. This study was a survey of H I in the galaxy. H I is an atomic gas very easy to detect in radio astronomy (21 cm of wavelength)
- the second one is an infrared survey using the IRAS[14] telescope, looking at the dust clouds

Why should they use these surveys? Dust clouds are made of atomic gas and molecular gas. If you know the dust distribution, you know the gas distribution. CO is well represented in molecular clouds. If you have the distribution of dust clouds[15] in our galaxy and you subtract from that the atomic gas,[16] you will get a map of the molecular gas that you can consider as a predictive map of the CO in the galaxy. The

---

[12] ‘Independent’ means here that the survey was made by a different team.

[13] ‘Independent’ means here, once again, that the survey was made by different teams. We shall elaborate on the notion of independence in Section 7.7.1.

[14] IRAS stands for Infrared Astronomical Satellite. It’s the first ever space-based observatory to perform a survey of the entire sky at infrared wavelengths.

[15] Astrophysicists get the dust clouds data using the IRAS survey.

[16] Astrophysicists get the atomic gas data thanks to the radio survey on HI.

scientists compare their data to those. So in this case we have independent surveys, instrument with the same physical principle/instrument with independent physical principle.

If we examine the five arguments presented in that paper, it appears that four of them are clearly a matter of robustness analysis. Each argument corresponds to a new derivation and each derivation gives new information about the phenomenon under study. Logically, on the basis of their own argumentation, the scientists conclude that results are robust, that is to say the maps represent only "reliable data", and the maps are complete. This first quick analysis has been done to illustrate, through this particular example that seems to us representative of many others, that robustness analysis is indeed a widely used, effective guiding principle of argumentation in scientific papers. Scientific papers are often a sequence of arguments organized according to a robustness scheme.

7.5 The Role of the Images in Scientific Papers' Argumentation

The images in scientific papers have traditionally been considered as mere illustrations. It is the propositional content, in contrast, which has been considered as essential and self-sufficient. Philosophers in particular have often underestimated the place of the images in scientific activities and results, victims of their "language-using ethnocentrism" as William C. Wimsatt has called it in his paper "Taming the Dimensions – Visualisations in Science" (1990). We believe that we must reconsider the role of images and their epistemic value in scientific papers. We must rethink their function in the composition and presentation of the robustness of results. In fact, we shall see that they play "a central role in the structure and the organisation of the scientific text. They are in fact the core of the scientific text" (Jacobi 1985). By "image", we take into account all what is non-textual in a scientific publication. That includes pictures, maps, graphs, histograms and so on.

Can an image constitute an argument? An image is not an argument if we take into account only its content and its internal structure. To be an argument, an image needs a textual support. The text actualizes some of the predication's virtualities contained in the image. The scientist who argues establishes some constraints that are guidelines for the final interpretation made by the reader. If the images structure the argumentation of a scientific paper, the text is complementary in two senses:

- The text reviews all the processes used to produce the image (instruments, data acquisition and data reduction). This is the "relay function" as defined by Roland Barthes (1964).
- The text limits the sense/meaning of an image in the caption. This is the "anchorage function" (Barthes 1964).

Let us return now to the analysis of our astrophysics paper and focus on the role of the images in the argumentation.

7.6 The Role of Images in the Milky Way Paper

At this point of our paper, we wish to focus more precisely on the details of the argumentation. The Milky Way paper has 10 pages. We find 20 figures which are either maps or curves. Basically, when we read the paper, what do we notice? It is made of text and images. We can ask now, what are the exact functions of the images in the argumentation we've briefly presented in Section 7.4.

7.6.1 The Role of the Images in the Reliability Claim

7.6.1.1 First Argument: The Role of Images in the Analysis of the Instrumentation and Data Processing in Order to Justify Their Reliability

There are no images in this section except one, showing the spatial position in the sky of the 37 studies. This is purely informative. There are no images of instrumentation. If the telescopes present anything out of the ordinary, they would probably have shown it on a photograph or a scheme. But this is not the case here, the telescopes used are very common in radio astronomy. For the data processing, there is no need for images because these procedures are common in astrophysics. In this section, the only real problematic point is the processing of the noise. The scientists dedicate the next section to this point.

7.6.1.2 Second Argument: The Role of Images in Comparing FCRAO Data with the CfA Data

As we said, the general claim to support here is the reliability claim. However, as soon as the scientists adopt one more derivation – here the FCRAO data – this general claim entails a more specific claim. The comparison between FCRAO data and CfA data aims to show that the particular noise processing done by the CfA scientists hasn't distorted the information obtained from the celestial sources. But they immediately face a problem: how can they compare the huge quantities of data acquired with the two telescopes? The challenge here is to find a relevant way to be able to compare them. In order to convince their reader, the scientists decided to associate three images.

These three images with the associated text are the argument ("grounds" in Toulmin's words). The caption stabilizes the meaning, in accordance with the anchorage function defined by Barthes.

Figure 7.2 enables a comparison between:

- dissimilar images (a) and (b).
- similar images (b) and (c)

It's not an isolated image that builds the argument but the joint use of a number of images. Images are juxtaposed so that they can be seen together. This is an

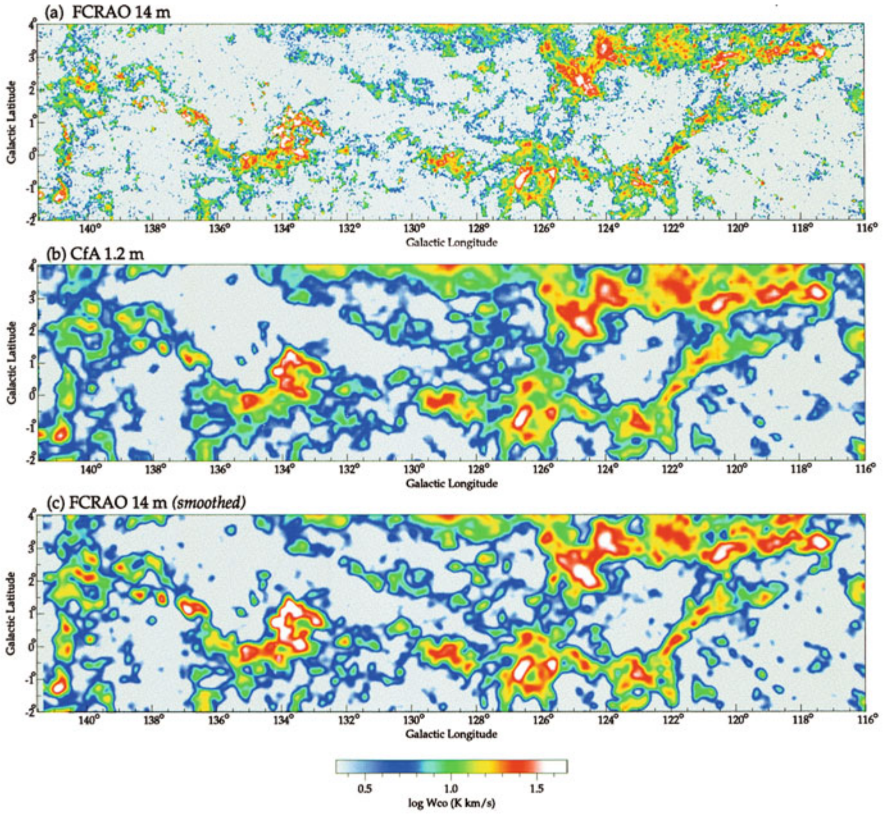

Fig. 7.2 (**a**) is the CO map of a specific celestial region. This map has been done by another team of researchers working with another radio telescope, the FCRAO, which has a ten time better resolution. (**b**) is the CO map of the same celestial region realized by the CfA radio telescope. (**c**) is the FCRAO map smoothed at the resolution of the CfA map. In accordance with its relay function, the main text describes the data processing from image (**a**) to image (**c**)
Source: Dame et al. (2001, CfA)

important point as emphasized by Tufte in his books (1990, 1997, 2007). Looking for differences and similarities requires this sort of comparative analysis. In our case, the argument consists in the visual similarities when one compares the images (b) and (c). Among the representational constraints, we have those relative to the best visualization. Using images enables:

- a better grasp of a huge number of data at the same time. It is necessary because of the overwhelming amount of data generated by the instruments.
- a better visualization of similarities or differences.

As one of the astrophysicists said about these images:

> This was my very first question about whether we analysed the data properly. (. . .) So, here is a comparison of the same region observed with two telescopes, independent telescopes.

> What I wanted to show is: everything that is on the map is real. (...) And then, you can see, amazingly well, these maps agreed. Even very small things, it's extraordinary. Better than I thought actually it would be. So, this was mainly to convince people that we've done everything right, because this is clearly independent data analysis. When you see that map, you believe it.

It would be unfeasible to compare directly the two sets of data produced by the two telescopes if they remain in numerical form. No one has the cognitive capacity to hold all items in the list in short-term memory, and then to do the calculations needed to extract conclusions about clouds' spatial localization and velocity. Taking some particular points on the map wouldn't be suitable either: in this case, one can always ask, what about the next point? Is it reliable? It would also be unfeasible to convert the relevant informational content of the map into a corresponding propositional content. Each map contains a huge number (potentially infinite) of predication's virtualities: there are the characteristics of each point and the links of each point with the others. Kitcher has stressed the same point about the Manhattan map (2001, p. 58):

> (...) the map is equivalent to a truly enormous number of claims about spatial relations: a picture is not worth a thousand words, but rather a staggering infinity of sentences. Further, although the map says many things that are incorrect, it also expresses an infinite number of true statements, for there are infinitely many truths of the form 'A is within Φ of being θ from due North of B', where A, B are places on the Manhattan shore and Φ, θ are angular measurements.

The fact that images can be potentially converted into a list of numerical data or into a propositional content doesn't mean that they are effectively converted in this kind of article. This is related to our limited cognitive abilities.

What is the strength of the astrophysicists' second argument?

If someone doesn't agree with the transition from these grounds (here the maps) to the claim (the reliability claim), he could ask for the warrant: the warrant here is that the different maps have been generated with two different radio telescopes in normal conditions of use. We find here a perfect example of the robustness' scheme in the sense of Wimsatt. This type of data processing is common in radio astronomy. If someone is not yet convinced by the warrant, he could ask for other general information to back up his trust in this particular warrant.

These telescopes and these procedures of data processing rely on well known physical theories. They are used in numerous studies without any problem. In the case of these three images, what is argued is: "Our data processing (and especially the processing concerning the noise) has not distorted the information you can see on the map".

7.6.1.3 Third Argument: The Role of Images in Comparing Optical Data with the CfA Data

Again, the general claim to support is the reliability claim. The comparison between the optical telescope and the CfA radio telescope aims to show that the

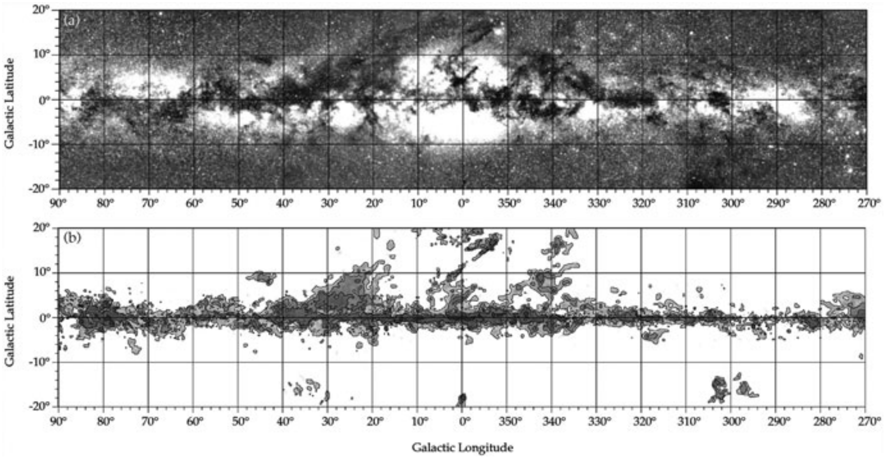

Fig. 7.3 (**a**) is a map created from 16 optical raw images of the galaxy. Like in radio astronomy, the procedure to put them together is called “mosaic”. (**b**) is the same map as Fig. 1a, but it’s zoomed on the center of the galaxy and put into grey scale
Source: Dame et al. (2001, CfA)

astrophysicists immediately meet the same problem: what can they do to compare the huge quantities of data acquired with the two telescopes in order to compare them effectively? They choose to create maps. In the Fig. 7.3, one can compare two reprocessed images.

The scientists zoomed into the center of the galaxy because it’s the brightest region in which emissions are at their highest level in optical wavelengths. The aim is to show that the optical light is obscured by dust clouds, so it is the best region to do that. Usually in astrophysics, to compare two images with two different wavelengths, the researchers place an image over another image using, for example, contours. They didn’t do that here. Why? Using contours was not a good way for representation, the visualization of the great similarities between these two images was not enhanced. Here is what one of the authors said:

> This was very challenging. It’s just a mass of dark clouds. I tried white contours, coloured contours, nothing worked because anything you put on top of this, because it is a great correlation, you’re getting CO emission where optical waves are. So I put a grid to help the eye, I couldn’t do it another way.

The grid here is very useful to understand how good the similarities are between these two maps. To enhance the strong correlation between the two maps, the scientists use grey scales. The association of the two maps together with the caption is the argument.

Claim: Our radio map is similar to the optical one. This argues that our map shows only reliable features (as we know that CO is present in dust clouds).

Grounds: To assert this claim, we rely on the following observational data:

- one map in optical wavelength of the center of the galaxy
- our radio map of the same region

Warrant: the different maps have been created by means of two different telescopes in normal conditions of use. This type of data processing is current in radio astronomy and in optical astronomy.

Backing: These types of telescopes and these procedures of data processing rely on well known physical theories. They are used in numerous studies without any problems.

7.6.2 *The Completeness Claim*

The last argument we will examine is the one that supports the "completeness claim": "Our survey of carbone monoxyde in the Milky Way is complete". Astrophysicists used several independent surveys:

- IRAS infrared survey who detected the dust clouds in our galaxy
- A radio survey of HI gas in our galaxy. HI is an atomic gas.

Why should they use these surveys? As we already described in Section 7.4.2, dust clouds are made of atomic gas and molecular gas. From dust distribution, you know the gas distribution. A way to obtain a predictive CO map is to subtract the atomic gas data from the dust clouds data (cf. Section 7.4.2).

The argument here takes place in two figures: Figs. 7.4 and 7.5. The basic procedure here is again to show similarities in the maps.

To enhance the predicted CO, they turn it in white on the map for a better visualisation you can see the details much better.

The argument goes on in Fig. 7.5.

If you look at this figure, you can compare easily and understand that there is a strong correlation between the two maps. Again, in this figure, the scientists tried to enhance the similarities. They chose the same colour scale for the maps, they chose to reprocess their maps to extract a profile to be even more persuasive.

7.7 Some Remarks on Independence and Invariance

Independence and invariance are two notions that turn out to be very important in the concept of robustness analysis. Indeed, without a precise definition of these two notions, the concept of robustness analysis loses its significance. What could we say about these two notions if we take into account our present analysis?

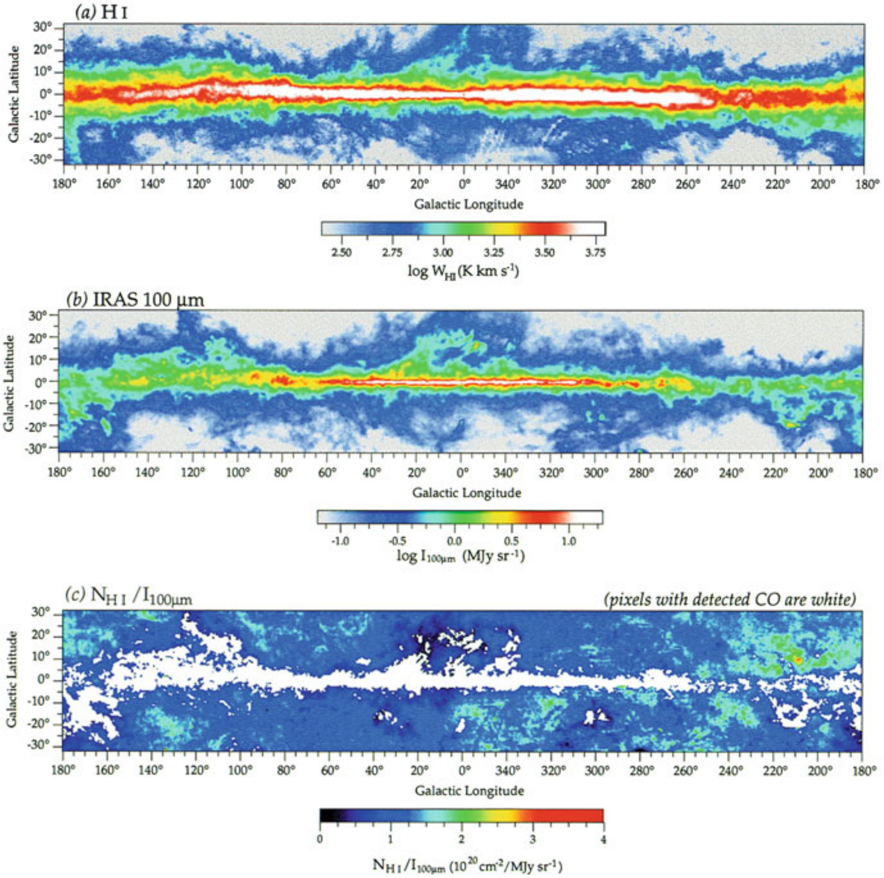

Fig. 7.4 (**a**) A radio map of the atomic gas (HI) in Milky Way done by another team with two other radio telescopes. (**b**) An infrared map of the dust clouds in the Milky Way done by another team with IRAS. (**c**) The IRAS map minus the atomic map (HI). It is in fact a CO predicted map *Source*: Dame et al. (2001, CfA)

7.7.1 Independence

Independence has to be understood in two different ways:

a) from a human perspective: From what the astrophysicist says in the interview, it is crucial that the data is produced by other teams of scientists. Why? Because scientists are aware of their own subjective choices. For example, in the noise treatment in our paper, it was crucial for the astrophysicists that the data processing didn't distort the data. So the fact that the radio data produced with

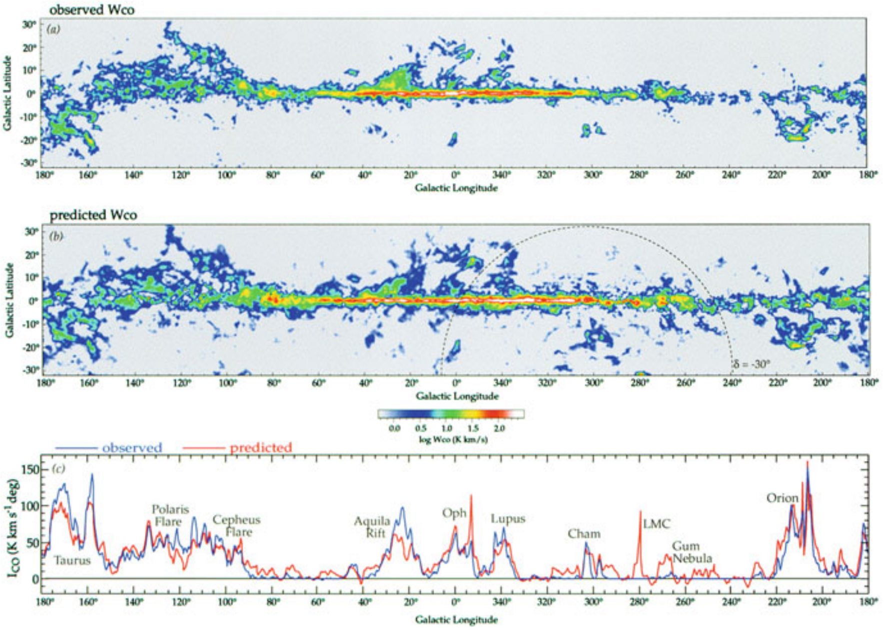

Fig. 7.5 (**a**) this is the CfA radio map. (**b**) the predicted CO map (the same that Fig. 4c but in another color scale). (**c**) is a more quantitative comparison. In *gray*: predicted CO and in *black*: observed CO. The *curves* are extracted from the two maps below
Source: Dame et al. (2001, CfA)

a better telescope (FCRAO) is similar allows them to be more confident in their own choices.

b) from an instrumental perspective: Astrophysicists use data produced by different types of telescopes (radio, infrared, optical telescopes). What does independence exactly mean when we talk about these different telescopes? In the case revealed in the Milky Way paper, telescopes record different wavelengths of the electromagnetic spectrum from radio wavelengths to gamma rays. One could say that the physical principles of these telescopes are not independent because they all allow the recording of electromagnetic radiations in the form of waves or particles. Each telescope allows the study of different properties of these molecular clouds: radio telescopes give information about the chemical properties of the molecular clouds; optical telescopes give information about the morphology of the molecular clouds, etc. In fact, the recording techniques and the technical problems are very different. These telescopes detect different properties of light (waves or particles). Each derivation (here different telescopes) enriches our understanding of the entity which is supposed to be the invariant under these multiple derivations.

7.7.2 Invariance

What is supposed to be invariant in our case? The localization and the velocity of the molecular clouds in the Milky Way, and more precisely, the localization and the velocity of CO which is located in these clouds. The spatial localization and the velocity are properties of an entity – molecular clouds – which are known by the detection of their properties. Different parts of the electromagnetic spectrum give us knowledge about different properties possessed by molecular clouds. So these clouds could be considered as entities which are clusters of properties. In fact, we could say that it is always properties and not objects that we primary observe – and not only in astrophysics. "Properties are primary, both metaphysically and epistemologically." (Humphreys 2006, p. 23). For Paul Humphreys, the ontological priority of properties suggests that the appropriate kind of realism to adopt here is property cluster realism. That kind of conceptualization seems to be a correct way to understand the way the scientists work. "The discovery of scientific entities thus involves a process rather like geographical discovery".

> (. . .) first an indistinct property is seen on the horizon; then some of the properties composing the coastline are discovered (. . .) then more and more details are filled in as further exploration takes place. We can be wrong about some of those properties – just look at the early maps of America, for example – (. . .) but we are rarely wrong about all of them. (*ibid.*, p. 25)

We have neglected a central point: the content of any image of the studied paper can be seen as determined by the causal relations involved in producing the data. Images cannot be understood only as symbols standing for something else. They are objects that have a causal relationship to the thing under study. Causal relations are relevant to understanding their role as evidence. Can the concept of robustness be fully developed without thinking about that point? This is a topic for another study.

7.8 Conclusion

As one can see, the argumentation in the astrophysics paper under study takes the form of a robustness analysis, in Wimsatt's sense: multiple derivations are mobilized to establish if the results are robust or not. In our example, the results are the two big maps. It is astonishing to consider as we did, at the same time, images as results and as arguments. In fact, they can have both functions depending on the role they play in the argumentation. For example, the two big maps of the article mentioned above are the results of this publication, and of course, as results, they must be, if possible, very robust. In order to increase this robustness, the scientists use a robustness analysis by exploiting other images published in other papers. But we can notice that if in the present argumentation these images are used as arguments, in the original papers where they come from, the same images were considered as results! One of the difficulties encountered by the researchers is precisely to find the best way to compare the data obtained with different detectors. This is done through data processing methods that can be extremely problematic. The arguments consist in the association of several images whose meaning, context and production procedures are stabilized by the text (caption or main text). Recognizing the similarities between the images lead to the conclusion that results independently produced converge.What the scientist aims to produce is a kind of Peircean cable constituted with many fibers (each fiber is an argument). A cable is robust because it is made of many fibers and, unlike a chain, the solidity of which depends on its weakest link, the cable remains robust even if one or two of its fibers break. (Callebaut 1993, p. 57).

To close this paper we would like to make some brief remarks about the relation between robustness analysis as we have characterized it through our example, and the realist/constructionist issue. If we remain at an epistemological level, our analysis could fit with the constructionist or the realist point of view. But if we consider the ontological level, it is certain that the astrophysicists use the robustness procedures in order to claim the existence of the entities/properties they studied. As philosophers, are we forced to endorse the point of view of the astrophysicists? Nothing is less certain. But nevertheless, as philosophers we have to take into consideration that the concept of robustness analysis (which is in fact a philosophical concept) gives a perfect account of the procedures used in the day-to-day activities of a lab to prove scientific assertions. That is what we wanted to show; the convergence here between the philosophical concept and concrete scientific practices. With the robustness concept, we have a perfect example of a "working" concept that can build a bridge between the scientists and the philosophers. We need that kind of "working" concept if we want to progress in our philosophical investigations of scientific practices.

References

Allamel-Raffin, C. 2004. "La production et les fonctions des images en physique des matériaux et en astrophysique." PhD diss., University of Strasbourg.

Allamel-Raffin, C. 2005. "De l'intersubjectivité à l'interinstrumentalité. L'exemple de la physique des surfaces." *Philosophia Scientiae* 9(1):3–31.

Barthes, R. 1964. *Eléments de Sémiologie*. Paris: Gonthier Médiations.

Callebaut, W. 1993. *Taking the Naturalistic Turn or How Real Philosophy of Science is Done*. Chicago: University of Chicago Press.

Chang, H. 2001. "Spirit Air, and the Quicksilver: The Search for the Real Temperature." *Historical Studies in the Physical and Biological Sciences* 31:260–84.

Collins, H.M. 1981. "Son of Seven Sexes: The Social Destruction of a Physical Phenomenon." *Social Studies of Science* 11(1):33–62.

Collins, H.M. 1992. *Changing Order. Replication and Induction in Scientific Practice*. London: Sage.

Collins, H.M., and T. Pinch. 1993. *The Golem*. Cambridge: Cambridge University Press.

Culp, S. 1995. "Objectivity in Experimental Inquiry: Breaking Data-Technique Circles." *Philosophy of Science* 62:430–50.

Dame, T.M., D. Hartmann, and P. Thaddeus. 2001. "The Milky Way in Molecular Clouds: A New Complete CO Survey." *The Astrophysical Journal* 547:792–813.

Daston, L., and P. Galison. 1992. "The Image of Objectivity." *Representation* 40:81–128.

Daston, L., and P. Galison. 2007. *Objectivity*. New York: Zone Books.

Godin, B., and Y. Gingras. 2002. "The Experimenter's Regress: From Skepticism to Argumentation." *Studies in History and Philosophy of Science* 3(1):137–52.

Hacking, I. 1981. "Do We See Through a Microscope?" *Pacific Philosophical Quarterly* 62:305–22.

Hacking, I. 1983. *Representing and Intervening*. Cambridge: Cambridge University Press.

Hacking, I. 1992. "The Self-Vindication of the Laboratory Sciences." In *Science as Practice and Culture*, dir., A. Pickering, 29–64. Chicago: University of Chicago Press.

Heyer, M.H., C. Brunt, R.L. Snell, J.E. Howe, F.P. Schloerb, and J.M. Carpenter. 1998. "The Five College Radio Astronomy Observatory CO Survey of the Outer Galaxy." *The Astrophysical Journal* 115:241–58.

Hudson, R.G. 1999. "Mesosomes: A Study in the Nature of Experimental Reasoning." *Philosophy of Science* 66:289–309.

Humphreys, P. 2006. *Extending Ourselves: Computational Science, Empiricism and Scientific Method*. Oxford: Oxford University Press.

Jacobi, D. 1985. "La visualisation des concepts dans la vulgarisation scientifique." *Culture Technique* 14:152–63.

Janik, A., R. Rieke, and S.E. Toulmin. 1984. *An Introduction to Reasoning*. New York: Macmillan.

Kitcher, P. 2001. *Science, Truth and Democracy*. Oxford: Oxford University Press.

Kosso, P. 1989. *Observation and Observability in the Physical Sciences*. Dordrecht: Kluwer.

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.

Putnam, H. 2003. "Pragmatisme et Connaissance Scientifique." In *Cent ans de Philosophie Américaine*, edited by J.-P. Cometti and C. Tiercelin, 135–55. Pau: Pup.

Toulmin, S.E. 1958. *The Uses of Argument*. Cambridge: Cambridge University Press.

Tufte, E.R. 1990. *Envisioning Information*. Cheshire, CT: Graphic Press.

Tufte, E.R. 1997. *Visual Explanations. Images and Quantities, Evidence and Narrative*. Cheshire, CT: Graphic Press.

Tufte, E.R. 2007. *The Visual Display of Quantitative Information*. Cheshire, CT: Graphic Press.

Waismann, F. 1945. "Verifiability." *Proceedings of the Aristotelian Society* 19:119–50.

Wimsatt, W.C. 1981. "Robustness, Reliability and Overdetermination." In *Scientific Inquiry and the Social Sciences*, edited by M.B. Brewer and B.E. Collins, 124–63. San Francisco, CA: Jossey-Bass Publishers.

Wimsatt, W.C. 1990. "Taming the Dimensions – Visualisations in Science." *PSA* 2:111–38.

Wimsatt, W.C. 2007. *Re-engineering Philosophy for Limited Beings*. Cambridge: Harvard University Press.

Winch, P. 2007. *The Idea of a Social Science and its Relation to Philosophy*. London: Routledge.