

Causal Order and Kinds of Robustness

Arnon Levy

The Hebrew University of Jerusalem

Forthcoming in "*Landscapes of Collectivity in the Life Sciences*" (MIT Press), Edited by Snait Gissis, Ehud Lamm and Ayelet Shavit

Introduction

Many biological systems are robust, in the sense that they retain their functionality in the face of external or internal perturbations. With the recent shift toward a systems-level perspective that has occurred in many parts of biology, the notion of robustness – which is essentially a system-level property – and the underlying processes that give rise to it, have received increasing attention (Kitano 2004; Sowyer et al. 2012; Wagner 2012). So have studies concerning the relationship between robustness and other system-level features, such as modularity (Alon 2007) and evolvability (Felix and Wagner 2008; Wagner 2008). My aim in this chapter is to offer a framework, or taxonomy, for thinking about different ways in which organisms and their parts achieve robustness. The framework is based on a notion of causal order, on which I have written elsewhere (Levy, 2014), and which appears to me a useful and important way to understand internal organization in biological (and perhaps other) systems. In a nutshell, orderliness is the degree of causal integration within a system (more details will be given below). I will argue, with the aid of examples and illustrations, that some forms of robustness stem from internal integration whereas others, by contrast, require a lack of integration.

There is an indirect but important connection between this way of looking at robustness, on the one hand, and the individual-collective relationship, on the other hand. Briefly, it can be put as follows. One

key difference between collectives and individuals appears to involve internal order and integration. Individuals are typically seen as integrated collectives, systems in which the orchestrated interactions among parts are constitutive of the behavior and functioning of the whole, often giving rise to a ‘greater than the sum of the parts’ situation. Collectives, on the other hand, typically comprise looser aggregates – the system is a “mere sum of its parts”. There is a clear sense in which the parts of such an aggregate belong to the same system, but their inter-dependence – and the dependence of the whole on their interactions – is weaker than in the case of individuals.¹ There is therefore reason to believe that, by understanding robustness in terms of order and disorder, we may arrive at an improved understanding of the functional similarities and differences between collectives and individuals. I will comment on this further toward the end of the paper.

Two short initial clarifications. First, my interest here is in robustness as a property of biological systems – or causal systems more generally – roughly, as stated, the property of retaining function in the face of perturbations. The term ‘robustness’ is often used to denote an epistemic property too – roughly, the accessibility of certain facts via different methods of discovery, measurement or justification.² There may be interesting connections between these ontological and epistemic notions of robustness, but it is the former that I will be discussing: my topic is robustness of the biologist’s objects of study, not of her knowledge or representations of them.

Second, as I understand the notion, robustness isn’t mere stability. A rock resting on the ground for a long period of time without moving, despite changes in its environment, is stable, but not robust. A robust system is one in which the stable feature is functional.³ One may also ask whether the feature’s robustness

¹ For a somewhat different approach to this contrast, see Wimsatt (1986).

² Woodward (2006) discusses different notions of epistemic robustness and the relations between them.

³ In biological systems such functionality can be understood, I will take it, in terms of the selected effects, although nothing in what follows depends on this.

is itself functional – whether robustness of this or that trait is adaptive, for instance. This is an issue that goes beyond the present discussion and I do not make assumptions on this score. I discuss how robustness is achieved but do not make assumptions about how or why robustness contributes to fitness or to some overall activity of the organism.

Causal order

Let me begin with a short characterization of the notion of causal order. Speaking generally, it concerns a distinction between kinds of causal systems or mechanisms⁴ – two ways in which a system's overall behavior may depend on the activities of its parts. It is useful, I think, to begin with a pair of contrasting examples. Consider the difference between, on the one hand, a macromolecular complex, such as the ribosome, and, on the other hand, a flux of diffusing particles. Both systems are constituted by many active parts, and these parts (and their activities) jointly determine the behavior of the whole. However, in the ribosome there is special significance to distinctions between parts and to the specifics of the system's internal layout in space and time. Each part of the ribosome plays a distinctive role and it is situated in the system in a particular way. That is to say: for the system to behave in as it does – here, the functional behavior of the ribosome, i.e. the production of proteins on the basis of mRNA – each part must act at a specific time and place, interact with specific components and not with others, and so on. In contrast, in a diffusive flux we see little of this internal heterogeneity: components (i.e. particles) are all but identical, and there are minimal constraints on the relations between them – particles are scattered about in a non-specific way, moving randomly, such that particular differences between them are unimportant. In this

⁴ Here I use the term 'mechanism' in a rough way to speak of causal systems in which overall behavior arises from the behavior of parts and their interrelations. For the connection between different notions of mechanism and philosophical projects associated with them, see Levy (2013).

kind of case, in contrast with the ribosome, the system behaves as it does due to the *lack* of distinctions and specific interrelations among underlying parts.

Intuitively, the ribosome is a very orderly system whereas the diffusive flux is very disorderly. I think this intuition captures something important that we can make more precise by way of the following definition.

Suppose we have a system S, exhibiting a behavior B (in the above examples, the systems are the ribosome and the flux, and the behaviors are protein synthesis and diffusion, respectively). We may say that S is orderly to the extent that:

(a) Distinct components of S play different roles in bringing about B.

And

(b) Components play their roles in virtue of local⁵ relations to other components.

The key idea is that in an orderly system there is a division of casual labor, so to speak, in which different parts play different roles and overall behavior is a function of these specific behaviors and interrelations. In contrast, in a disorderly system overall behavior is a cumulative outcome of the behavior of multiple (often many) similarly acting parts. The specifics of parts and interrelations are unimportant.

This definition can be clarified and elaborated considerably (see Levy 2014). But for present purposes I will make only two remarks: First, order and disorder, as I use these terms, are always relative to some focal behavior (this is embodied in the above definition). For instance, the ribosome produces proteins in an orderly fashion, but it dissipates heat in a disorderly way. In what follows the focal behavior, i.e. the feature I will concentrate on, is the system's robustness – its ability to maintain function in the face of

⁵ 'Local' may denote spatial proximity, but other sorts of 'closeness,' such as closeness within a causal network, may be relevant too.

perturbations. It is this feature that will be assessed in terms of its orderliness and such an assessment does not carry implications for the orderliness of other features of the system. Second, some phrasings below may give the impression that I treat order and disorder as discontinuous, on/off categories. But this is just a convenient way of speaking. These are graded categories, matters of degree, and talking in terms of order *versus* disorder is merely a way of simplifying the presentation.

Orderly Robustness

With these ideas introduced, if briefly, let me now describe three types of robust mechanism: orderly, disorderly, and a hybrid or semi-orderly type. Each has interesting sub-types too, some of which I'll describe.

Orderly robustness is robustness that's due to an internal division of labor: differences and interrelations between components are responsible for resilience to perturbations. A simple but important form of orderly robustness is what's often called *functional redundancy*. A functionally redundant system has more than one instance of a certain component or process, so that if the component fails, the system detects this, and activates a functionally identical backup component.⁶ This sort of process is orderly because there must be a clear distinction between the "ordinary", default components and the backup, as well as a mechanism that detects situations in which the default component fails such that the backup needs to be activated. These are distinct and the relations between them are vital to the system's robustness. Functional

⁶ This formulation is intentionally abstract, specifically with regards to 'detection.' The idea is that the system is set up such that if the default component fails it has the capacity to adjust its behavior and activate the backup. Such detection may reside in a designated component, a detector as it were. The detector may be sensitive to the activity of the default component *per se*, or to the system's level of output. These differences may matter in some contexts, but I will not explore that here. For present purposes, what matters is only that there is a distinction between the default/backup components and some other factor that 'decides' which one to activate.

redundancy is very important and prevalent. But since it is conceptually straightforward, and since there is an extensive discussion of empirical examples, especially in gene regulation, I will not discuss it further.

A second kind of orderly robustness is more subtle, and perhaps less familiar. Here the system is set up so that a fault in one component is corrected by one or more *different* components. Unlike redundancy cases, there need not be duplication or some other form of backup. Instead, the system retains functionality by adjusting the level or order of activity in some of its sub-parts, ones that would have been “on” anyway, so as to compensate for the perturbation. Let us call this *checks-and-balances robustness*, since it is a case where different components of the system correct the output by performing different roles (and not via backup). Further distinctions can be made within this category, and I will mention two important ones. One kind of case that can be placed in the checks and balances category consists of a system, such as a cell, that is able to divert some or all of the activity of a given causal (e.g. biochemical) pathway to another pathway, producing an equivalent output. This is somewhat similar to redundancy, but the backup process need not be identical to the default one. For instance, Wagner (2005) discusses the Pentose-Phosphate pathway in *E. coli*, which metabolizes glucose. A failure in one of the components of this pathway, the enzyme *glucose 6-phosphate dehydrogenase*, turns out to be compensated for by increased activity of the TCA cycle. The TCA cycle breaks down glucose as well, but does so in a different way. So it is a case of checks and balances, as defined above. Note that this form of robustness will often be cheaper to maintain, in fitness terms, relative to functional redundancy. For it does not involve maintaining extra, redundant copies. Instead it regulates the activity of already-on components so as to maintain a stable level of output. But there is a flip side: the “failsafe” operation has a different set of side effects. For, unlike redundancy, the regular process is *not* identical to the process as it works under perturbative conditions. Therefore, it may have substantially different side effects. This is indeed what occurs in the case of glucose metabolism in *E. coli*, in which the shift from the Pentose-Phosphate pathway to the TCA cycle is accompanied by a shift in the balance of NADPH production.

Within the checks and balances category I think we may also include integral feedback, a process that is well understood in engineering and that has been found to be present in several important biological contexts (Sovanshi et al. 2015). Briefly described: In integral feedback the system is sensitive to the cumulative “error” relative to some target output, and adjusts its behavior accordingly – typically, the value of some variable is adjusted upwards or downwards. The simplest way to achieve this is via negative feedback, as in a simple thermostat. A more important (and also more complex) case from biology occurs in bacterial chemotaxis, modeled by Barkai and Leibler (1997). Here, the bacterium exhibits so-called sensory adaptation given changes in the concentration of a ligand in its environment. It retains the ability to respond to the ligand’s gradient despite changes in the ligand’s concentration. This response, Leibler and Barkai show, is robust to changes in the system’s internal composition, including dramatic changes in concentration. Such a mechanism of integral feedback, it will be noted, is especially relevant where the system must be robust against fluctuating inputs or noisy concentrations of components. The notion (and mechanism) of integral feedback is applicable only when there is a well-defined, continuously adjustable system-level parameter.

A number of recent theoretical studies have looked into relatives and generalizations of integral feedback. For instance, Hart and Alon (2014) review biological networks that contain what they dub *paradoxical* components, namely a component that exerts both excitatory and inhibitory influences at once. They show it can endow a network with robustness to fluctuating concentrations. In a still more general study, Shinar and Feinberg (2011) derive explicit conditions for the kinds of networks in which one may expect certain forms of robustness. These conditions depend on subtle features of the network’s topology, especially on the network’s so-called *deficiency* – a quantity that, roughly speaking, serves as a measure of the interdependency between sub-parts of the network.

Taken together, these and other studies suggest that orderly, integral feedback-like mechanisms may well be quite common and important, and that they take quite diverse forms. Moreover, there may well be theoretical principles governing the connection between a network's setup and its degree of robustness.

A final kind of orderly robustness I will mention is what might be called *quality control* robustness. In this type of mechanism, an error-prone underlying process (typically an orderly one as in the example below) is monitored by a separate orderly process, which corrects the former's errors. Here too there is an important role for a division of labor within the system, and so it is orderly in the sense described earlier. However, it is neither a case of redundancy nor an integral feedback like scheme. DNA proofreading is a key example in this category – the underlying process is DNA replication, involving many distinct parts and steps, and its results are monitored by a separate mechanism, which effectively compares the replicate to the original DNA sequence, correcting replication errors. There are other examples of a broadly similar sort. It is helpful, I think, to see these cases as a sub-class of orderly robustness.

Disorderly Robustness

Let's now look at the other end of the spectrum, at disorderly robustness, i.e. robustness without an internal division of labor. Speaking generally, disorderly robustness is achieved via aggregating the effects of multiple, similarly acting components, such that system behavior does not depend on any one (or small number) of them.

A simple form this might take involves a kind of redundancy, but not functional redundancy in the sense discussed above. Let me start with a non-biological example. Consider the construction of massive manmade objects, such as bridges and buildings. Typically, such structures have substantially more load bearing beams than are needed to carry the mass of the structure in normal circumstances. This is a design

feature that enables them to be robust against various insults – for instance, in the face of a fracture in one of the beams. In this kind of case, all the components (load bearing beams) are acting simultaneously and in the same way, but there are more beams than needed to support the load. We can call this *overkill* robustness. Overkill robustness is somewhat wasteful, since it involves spare capacity. But it is developmentally (relatively) simple and one can readily see how it might have evolved. I am not confident as to how common such robustness is in biological systems, but there is at least one significant case of it, namely kidney capacity. Most people can maintain normal renal functioning even if one of their kidneys fails or gets removed. In part, that is because having two fully operational kidneys is, in normal circumstances, overkill.⁷

A second kind of disorderly robustness operates via a different sort of aggregation. In a classic paper entitled ‘Physics of Chemoreception,’ Howard Berg and Ed Purcell (1977) ask the following question: Suppose a cell must detect the level of some chemical signal that is diffusing in its extra-cellular medium. It does so via receptors on its surface that absorb and quickly clear the ligand. How much of the cell's surface must be covered by receptors for the cell to efficiently detect the signal? Berg and Purcell show that, given certain mild idealizations, a cell in which far less than 1% of surface area is covered by receptors does as well in terms of detection capacity as a cell in which the whole surface is packed with receptors. This surprising result is due to the fact that a diffusing ligand in effect performs an extensive random search on the cell surface. This allows the cell to reduce, so to speak, the concentration of receptors on its surface by orders of magnitude. Berg and Purcell do not use the term robustness, but it is possible to view their result in this fashion. The cell does as well with 0.5% receptor coverage as it does with 50% – it is robust to (large) changes in the concentration of receptors. But this is done in a disorderly

⁷ When a kidney is removed the other kidney will often expand somewhat, and nephronic tubes within it become enlarged. But, as far as I've been able to find out, this doesn't bring the single kidney back to the filtering capacity of two healthy ones; it's just that we're born with "overabundant... kidney capacity" (Andres 2002).

fashion: all receptor molecules are identical and no particular organization is required. Such a process does not involve functional redundancy, since none of the receptors serve as backup. Nor is it a form of overkill, as the receptors do not operate simultaneously. It is a different form of robust process: a kind of trial and error, done on a massive scale.

Semi-Orderly Robustness

The final category I wish to describe, rather briefly, combines the orderly and the disorderly. A paradigmatic example is that of synapse elimination. In several neural developmental contexts, there is a process whereby a large number of synapses is formed – for instance, between the central nervous system and a peripheral muscle – but only a subset of them is retained into adulthood. To simplify considerably: synaptic connections that receive positive postsynaptic feedback are kept, while others atrophy. The process involves a form of selective retention: only synapses that do a “good job” are maintained. In consequence, the neuromuscular junction develops the right level and quality of enervation, despite substantial variation in both internal and environmental stimuli. Hence the development of the neuromuscular junctions is robust.⁸

Selective retention can be thought of as a semi-orderly process. On the one hand, it resembles disorderly robustness since many similar elements act at once, at least in the initial phase of the process. On the other hand, it is similar to an orderly mechanism, like quality control, because there’s an element of inspection, a weeding out of undesirable results. So I think we can understand it to be a hybrid or a superposition of the

⁸ Let me be clear that I am not claiming that the central role of the mechanism of synapse elimination is to ensure robust neural development; or at any rate, that is not a claim that is required in the present discussion. I am only pointing out that synapse elimination makes a robust mechanism relative to other possible ways of, for instance, forming neuromuscular junctions.

orderly and the disorderly. It utilizes the search-power of disorderly aggregates, but matches a pre-specified target via an orderly, division-of-labor type process.

Now, there may be other semi-orderly types of mechanisms, but this is the only one I have been able to pinpoint as biologically realistic. I think it is helpful as an illustration of a potentially more general category, and as a contrast to the types of robustness discussed earlier.

Order, disorder and individuality

As I noted early on, I think the order-disorder axis marks a key dimension of difference between individuals and collectives. In closing, let me elaborate somewhat. It is common to think of an individual – in contrast to a collective – as an integrated unit. To be sure, such a characterization does not capture any specific type of individual, but it may be seen as a baseline – an organism is an integrated unit with certain other features (metabolism? The ability to reproduce? An immune system?). My tentative suggestion is that the notion of order I have outlined captures much of what is meant when a system is regarded as causally integrated: It has diverse parts and these parts interact in ways that depend on the identity of the parts and their specific organizational features. If this kind of framework for thinking about individuals is accepted, and given how the varieties of robustness I've discussed relate to it, some insights and further questions can be outlined.

A first point to note is this. I have given examples primarily from cellular and molecular biology. Or at any rate from intra-bodily phenomena. But the definition of order applies to any complex system in which interactions among parts determine the behavior of the whole. In particular, it applies to collectives that are made up of internally ordered systems, such as organisms. Such collectives of individuals may themselves function as they do due to a more orderly underlying structure, as is the case with many of the features of

social insect communities. But the underlying dynamics of a collective made up of individuals need not be orderly. It may operate in a manner closer to a diffusing flux or a buffalo herd.

Given the taxonomy I have offered, such broad structural analogies may also allow us to think about the ways in which such collective systems may be robust, and perhaps to gain some understanding of the reasons why they are structured as they are. Consider a flock of birds or a school of fish moving as a collective. Current work on collective motion invariably regards these phenomena as disorderly processes, in which individuals (birds, fish) move in a simple, largely identical manner, according to a small number of simple behavioral rules. This is often described by stating that such behavior is *self-organizing*,⁹ i.e. that it stems from the interactions of a large number of autonomous individual organisms, without centralized control and direction and without a pre-existing template.¹⁰ For the most part, existing investigations pertain to the proximate mechanisms at play, and to how they result in the collective behavior at the ‘first order,’ as it were. But now consider the question of robustness (a question that, as far as I know, has received very limited attention in this area). It seems unlikely that a flock's behavior can be adaptive unless it is robust – especially to idiosyncrasies in the behavior of individuals and/or to injury or departure of an individual from the flock.¹¹ If this is so, we may suggest an evolutionary rationale for the fact that collective motion is typically self-organized, rather than being a more orderly, centralized behavior (as in cases of collective hunting, for instance). Self-organization allows the flock to be robust in a way that may not be available, or that would be expensive in terms of the fitness of individual members of the collective,

⁹ As Ehud Lamm has pointed out to me, it is possible to treat self-organization as itself a kind of robustness – e.g. robustness to the existence of a leader, in the case of collective motion. But in order to keep the presentation simple I will not discuss it in this way.

¹⁰ Some models – individual-based models – assume a set of simple rules that each individual executes, and then demonstrate that a large collection of such individuals behaves as does a flock or school (Lopez et al., 2012). Other models operate more abstractly, treating the collective as a physical aggregate and using methods from statistical physics (Vicsek and Zafeiris 2012). The differences between these approaches are interesting, but do not affect the present discussion.

¹¹ This is especially plausible on the assumption that many flocking and schooling behaviors evolved, in large part, as a means for predator avoidance (Parrish and Edelman-Keshet 1999).

if a single, or small number of individuals served as group organizers, or if a pattern had to be "memorized" via some special arrangement in the flock. This suggestion is, of course, merely a suggestion, a broad-brushstroke hypothesis. But it serves to illustrate how thinking in terms of the categories I've described may help expand our understanding of collective phenomena: The causal organization of a system such as a flock is disorderly in a fairly radical way. When wondering why this is the case we may point, as one possible reason, to the kind of robustness it achieves.

Another potential set of lessons concerns collectivity within an individual. While it is commonly accepted that collectives are made up of individuals, the foregoing discussion illustrates that individuals may, in one sense, be composed of collectives. Or more accurately, some functions within individuals may be carried out by collectives. Such is the case with disorderly systems within an organism, and to a certain extent with semi-orderly ones (sections 4 and 5). My discussion illustrates this through functional systems that occur within an individual but are akin to a collective in terms of their causal setup, such as Berg and Purcell's receptors or, to an extent, synapse elimination.

But beyond the conceptual interest in this kind of possibility, the foregoing discussion suggests one reason why we should expect such a situation – namely, disorderly systems permit a certain kind of robustness. If this is correct then it would be important to better understand the specific advantages conferred by attaining robustness in a disorderly fashion, and the contexts in which we may expect such a collective-within-an-individual situation. Do orderly and messy robustness have different fitness profiles, such that they are selected in part due to the fitness advantages that they confer? Or, on the contrary, is their selection primarily on the function, the effect itself, with robustness severing as an 'added bonus'? Are there differences in terms of evolvability between different kinds of robustness – is a messy system more readily adaptable, at least to some conditions than an orderly one? Are disorderly and/or semi-orderly systems forms of robustness more prevalent in some developmental contexts than in others (e.g.

early relative to late differentiation), and if so, why? I think these questions can benefit from being cast in the light of the order/disorder distinction, because that distinction tells us something about how the system achieves its function, and this gives us clues about what advantages and drawbacks it has in other respects (evolutionary, developmental etc.). Here I can only raise such questions, however, and in a preliminary form at that. My hope is that they may stimulate further elaboration and investigation.

Summary

Let me summarize briefly. I have described three broad forms of achieving robustness: orderly, disorderly and semi-orderly. Orderly robustness arises from an internal division of labor, either of the backup sort, or of a more sophisticated kind as in integral feedback and quality control. Disorderly robustness arises from the cumulative effects of a homogenous collection of elements. In simple cases this involves overkill, i.e. the existence of spare capacity. In more complex cases there is a form of massive trial and error. In between, we find semi-orderly processes, like selective retention, in which there is a messy phase or aspect, on top of which there operates an orderly process that “reins in” the mess. The examples I gave were primarily from cellular and molecular phenomena, but I think the general framework is relevant to other types of processes too. Finally, I connected this discussion, albeit tentatively, to one dimension of individuality – namely the degree of causal integration – and suggested that we may be able to learn interesting lessons about the relations between individuals and collectives at different scales by looking at them through the lens of order and disorder.

References

Alon, U. 2007. *An Introduction to Systems Biology*, Boca Raton, FL: CRC Press.

Andres, M.A.W. (2002), "How Can You Live without One of Your Kidneys?" *Scientific American online*: <http://www.scientificamerican.com/article/how-can-you-live-without/>.

Barkai, N. and S. Leibler. 1997. "Robustness in Simple Biochemical Networks." *Nature*, 387: 914-917.

Berg, H. and E. Purcell. 1977. "Physics of Chemoreception." *Biophysical Journal*. 20(2): 193-219.

Félix, M.A. and A. Wagner. 2008. "Robustness and evolution: concepts, insights and challenges from a developmental model system." *Heredity*, 100(2): 132-140.

Hart, Y. and U. Alon. 2013. "The Utility of Paradoxical Components in Biological Circuits." *Molecular Cell* 49(2): 213-221.

Kitano, H. 2004. "Biological Robustness." *Nature Reviews Genetics* 5: 826-837.

Levy, A. 2013. "Three kinds of New Mechanism." *Biology and Philosophy* 28(1): 99-114.

Levy, A. 2014. "Machine-Likeness and Explanation by Decomposition." *Philosopher's Imprint* 14(6): 1-15.

Lopez, U., J. Guatrais, I.D. Couzin and G. Theraulaz. 2012. "From behavioural analyses to models of collective motion in fish schools." *Interface Focus* 2: 693-707.

Parrish, J. K. and L. Edelstein-Keshet. 1999. "Complexity, pattern, and evolutionary trade-offs in animal aggregation." *Science* 284(5411): 99-101.

Shinar, G. and M. Feinberg. 2011. "Design principles for robust biochemical reaction networks: What works, what cannot work, and what might almost work." *Mathematical Biosciences* 231(1): 39-48.

Sovanshi P. R., A.K. Patel, S. Bhartiya, K.V. Venkatesh 2015. "Implementation of integral feedback control in biological systems." *WIREs Systems Biology and Medicine*. 7(5): 301-316.

Sowyer, A. (Ed.) 2012. *Evolutionary Systems Biology*. Berlin: Springer.

Vicsek, T. and A. Zafeiris. 2012. "Collective motion." *Physics Reports* 517(3-4): 71–140.

Wagner, A. 2005. "Distributed robustness versus redundancy as causes of mutational robustness." *BioEssays* 27(2): 176-188.

Wagner, A. 2012. "The role of robustness in phenotypic adaptation and innovation." *Proceedings of the Royal Society B*. 279(1732): 1249-1258.

Wimsatt, W. 1986. "Forms of Aggregativity." in Donagan, A., A.N. Perovich and M.V. Wedin (eds.), *Human Nature and Natural Knowledge*. Dordrecht: Reidel.

Woodward, J. 2006. "Some varieties of robustness", *Journal of Economic Methodology*. 13(2): 219–240.

Acknowledgments: For discussion of a previous version of this paper I am grateful to Rosa Cao, Sandra Mitchell and Andreas Wagner. Ehud Lamm provided insightful editorial comments that improved the paper considerably.