The Dynamics of Risk-Taking:
How Damage Affects Real-World Complex Entities, and a Framework to Estimate the Impact of Policies and Technologies on Their Survival

Luca Dellanna

luca-dellanna.com

Luca@Luca-dellanna.com

November 27, 2019

last revised April 5, 2020

Abstract

"*Choose policies not for what they intend to do, but for how people react to them.*"

To be able to design policies, technologies and organizational structures that protect us from risk rather than exposing us to it after second-order effects are taken into account, we need to understand why people and organization take risks and how they become more or less vulnerable to them.

People and organization are complex entities, and damage has complex effects upon them. Some forms of damage kill them and other ones make them stronger. Some form of protection protect them, whereas others make them take more risk, eventually resulting in tragedies.

Classical behavioral economics often uses one or more assumptions not representative of the real world: static monolithic players playing single gambles single times with known static odds. **This paper discards these assumptions and instead considers complex players made of multiple levels of components (e.g., muscle fibers for humans or citizens for cities) with non-homogenous and anisotropic internal structures, exposed to multiple risks (social, financial, reproductive, etc), engaged in multiple repeated interactions with unknown dynamic odds and who undergo plastic changes in both their physique and in their behavior when damaged or rewarded.**

This paper adopts a bottom-up approach to build a coherent and comprehensive framework to estimate the long-term effect of policies, technologies and organizational structures on the risk-taking behavior of people and organization in the real world and on their ultimate survival. In particular, it focuses on how damage can both make them stronger or weaker, more or less prone to take risks, and on what it depends.

Contents

Introduction

The objective of this paper is to provide the reader with a set of tools he or she can use to evaluate the impact of policies, technologies and organizational structures on the risk-taking and on the survival of those affected by them.

It is a long paper – but making it shorter would have implied using fragile simplifications and provide top-down confabulations for why things work the way they do. Conversely, taking the time to understand the bottom-up processes that cause people and organizations to break and to adapt will allow the reader to build a much more useful and robust model.

i. Applications

A reader who is a manager will find this paper useful to gauge whether the policies and organizational structures of his company cause its employees to take excessive risks and will learn how to develop better ones that allow its company to rapidly adapt to changes in the market.

A reader who is an investor will find tools to gauge the fragility of companies and countries he is deciding whether to invest in.

A reader who is a policy-maker will learn how not to design policies with harmful and difficult-to-predict side-effects.

And regardless of his role, every reader will gain tools to understand some of the ways in which things, people, companies and societies break (and what to do about it, to take care of his or her loved ones).

ii. Real-world risk-taking

The following six assumptions, recurrent in classical behavioral economics, lead to models which do not represent what is going on in real-world risk-taking.

1. **Independent activities.** In classical behavioral economics, decision making regarding any activity is studied as if people were subject exclusively to the risks proper of that activity. However, observed phenomena such as risk homeostasis (in which, when given the choice between increased safety and increased performance, people choose the latter) do not square with this assumption.
   Risk homeostasis is explained by the fact that in real-life, people are exposed to multiple sources of risk, and generally take choices to maximize their survival across all of them – for example increasing risk-taking at an activity irrationally *from the point of view of survival at that activity* in order to increase its payoff and use the gained resources to manage other sources of risk (a rational behavior if survival *across all activities* is considered). Risk homeostasis will be discussed more comprehensively in section IV.
   This paper considers multiple intertwined activities and risks.

2. **Averageable outcomes.** In classical behavioral economics, gambles are studied assuming that the outcome of people playing them multiple time coincides with the outcome of multiple people playing them once. For example, it is assumed that the payoffs of of n people playing Russian Roulette once is the same of that of one person playing it n. This is not the case: for large n, the outcome of the former case is n times $\frac{5}{6}$ (the survival rate) times

the rewards for surviving, whereas the outcome of the latter is death. The relatively new field of ergodicity is tackling this problem, with scholars such as Ole Peters and Ed Thorp. This paper does not assume ergodicity.

3. **Known odds.** In classical behavioral economics, gamblers know their odds or can reliably estimate them (because their distribution is bounded). In real-life, instead, players often face uncertain odds and put themselves at risk if miscalculating them. Nassim Nicholas Taleb's *Fooled by Randomness* describes many of such phenomena.
   (Note: unknown odds are not the same to uncertain outcome from known odds, but an additional source of uncertainty.
   This paper considers uncertain odds.)
4. **Static odds.** Not only are the odds of most activities unknown, but they are also constantly evolving, due to the environment being an adaptive entity. For example, one of the reasons for why it is impossible to come up with a perfect strategy to maximize stock market gains is that, if such strategy existed, other investors would adapt to it, changing the way the market behaves and consequently changing the odds of the strategy.
   This paper considers dynamic odds.
5. **Static behavior.** In classical behavioral economics, a rational player would play a gamble the same way every time. In the real world, behavior is path-dependent: players often change their strategy based on its last few outcomes.
   This paper considers dynamic behavior.
6. **Static entities.** In classical behavioral economics, gambles are studied assuming that they only affect players in the measure that payoffs increase or decrease their (usually financial) resources. In real-life, people performing risky activities might undergo damage which plastically changes them. For example, a person undergoing exposure to small doses of a poison becomes more resistant to it, a process called *hormesis*. Nassim Nicholas Taleb's *antifragile* talks at length about such as humans are physically changed by damage, including the previous example. This paper will explain the bottom-up processes that cause antifragility, how the antifragile *can* express fragile behavior, and will describe how damage changes risk-taking behavior.
   This paper considers dynamic entities.

The six assumptions described above do not apply to the real world, and are thus not used here. This paper builds a coherent framework which describes the dynamic risk-taking behavior of people and organizations in the real world, with a focus on how damage (or lack thereof) affects the resistance and risk exposure of those affected by it.

iii. Damage as the blueprint for adaptation

People, organizations and, in general, entities that adapt to the environment have two basic survival needs: resources (the building block of their constituents), and damage (the blueprint for prioritizing use of the resources)[1].

The direction in which to adapt has to match the one required by the environment; otherwise, maladaptation takes place. Experiencing what can damage muscle fibers, for example, signals to

[1] In fact, it will be shown in section II that in case of a static risk profile from the environment, an adaptive entity will adapt in order to match the risk profile (the most efficient use of resources to maximize survival). In that sense, the environmental risk profile as measured through damage is a blueprint.

our body to direct nutrients towards building larger muscles; a sedentary lifestyle where muscle fibers never get damaged directs nutrients towards building fat and causes muscles to atrophy. **Damage is associated to feedback used by adaptive entities to determine what is required to survive the environment.**

Behavior is largely a reactive response to damage or to threats which caused damage in the past and might cause it again in the future. Hence the need to understand how damage affects the people and organizations before discussing their behavior. Damage will be the common thread of this paper.

Nor wisdom nor experience can be transferred from master to student; only knowledge. The difference is damage. Wisdom and experience are knowledge filtered through the lens of damage. Smart theories can break under the test of practice. Damage is needed to filter what is erroneous from what is correct.

The advantage of using damage as the blueprint for adaptation, rather than other alternatives, is that it doesn't require any assumption other than being exposed to the very environment the entity must survive into. It doesn't require any prediction[2], nor it relies on measurement of some proxies which might be correlated to survival today but might not be it tomorrow. Damage is always the right proxy to measure vulnerability – assuming only exposure to the right risk profile[3].

Hence, the importance for damage, the object of the first part of this paper.

iv. Outline

This paper is organized as follows.

The first section, "Damage Mechanics: How Things Break", describes how things, people and organizations incur damage and what kind of damage causes them to stop functioning as intended. The fundamental notion of *not all damage is equal, pattern matters* will be discussed.

The second section, "Adaptation Mechanics: How Entities Adapt", describes how things, people and organizations adapt to their environment. In particular, it explains what makes an entity antifragile and how different exposures to the environment causes it to be more or less likely to exhibit an antifragile response.

The third section, "Adaptation Dynamics: Tinkering With Adaptation", describes how changing the internal structure, exposure and adaptation rate of antifragile entities changes both their behavior and their likelihood of exhibiting an antifragile response. It will also discuss how convexity to damage changes over time.

The fourth section, "Exposure Dynamics: Why People Take Risks", describes how people adjust their exposure to risks in the real world, adopting *seemingly* irrational behaviors. It will also discuss redundancy over space, time and scales.

The fifth section, "Ergodicity: How to Reap The Gains of Antifragility", describes some principles to ensure survival and discusses them in the context of ergodicity and morals.

The sixth section, "Policies, Technologies and Organizational Structures: How to Avoid Fragilizing Ourselves", uses the framework assembled so far to describe the effects of policies, technologies and organizational structures on the survival of the people and communities who adopt them.

Each section contains examples, principles, justifications, limitations and practical applications.

[2] [Taleb, 2012] has a section on the perils of using predictions.
[3] What this assumption means will be explained in section II.

I. Damage Mechanics: How Things Break

Damage breaks things, people and organizations. However, not all damage is equal: understanding damage patterns and how damage propagates is the key to understand why some things are more resistant than others.

In this section, I'll describe a framework which describes how both living and non-living entities get damaged. This chapter takes concepts from materials science and generalizes them to describe how entities of any type break.

A note for the readers who already know the difference between fragile and antifragile entities: this section applies to both. Antifragile entities differ from fragile ones for how they react after they get damaged, not in how they get damaged.

i. Damage and stressor magnitude

Things break in two patterns: accumulated damage from multiple events and sudden damage from a single extreme event.

Imagine a man hitting a rock with a hammer. If he strikes it very lightly, he will cause no damage, no matter how many times he hits it. If he applies moderate strength, each hit might create some micro-cracks. For now, no damage is observable at the macro level and the rock still seems very sturdy. The damage is fully contained at the level of the constituents of the rock – the stone crystals – and only observable at that level. However, this micro-damage accumulates; if the man keeps hitting the rock, eventually, it will break. If, instead, the man begins hitting the rock with a huge hammer and superhuman strength, the stone might break in a single hit.

The same is true for living beings. If a man lifts a matchbox a hundred times, no damage will occur to his muscles. If he lifts a moderate weight a hundred times, he will create some micro-tears in his muscle fibers (and muscle growth will be triggered – but this is the topic of another section). This damage also accumulates, if repeated multiple times over short time, before healing can take place. If the man lifts a too-heavy weight, he might tear his muscle, causing an injury.

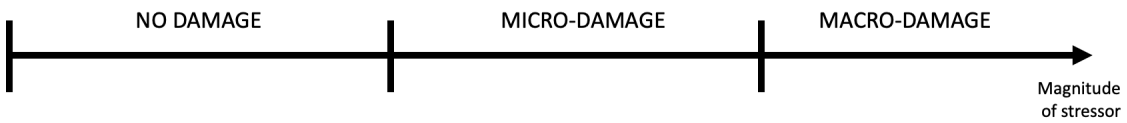

Figure 1: *A representation of damage as a partial function of stressor magnitude. The term "macro-damage" used here will be changed to the more correct term of "functional impairment" later on in this paper, once the homonymous concept will have been introduced.*

To understand the difference between micro-damage and macro-damage, and how much of the former constitutes the latter, it has to be understood that most entities are not monolithic, but are made of components, which might be made of components themselves. For example, a human muscle is made of muscle fibers which are themselves made of small cylindrical structures called myofibrils. As another example, species are made of animals and civilizations are made of cities or tribes which are made of families which are made of people.

These components are organized in levels, as exemplified in fig. 2.

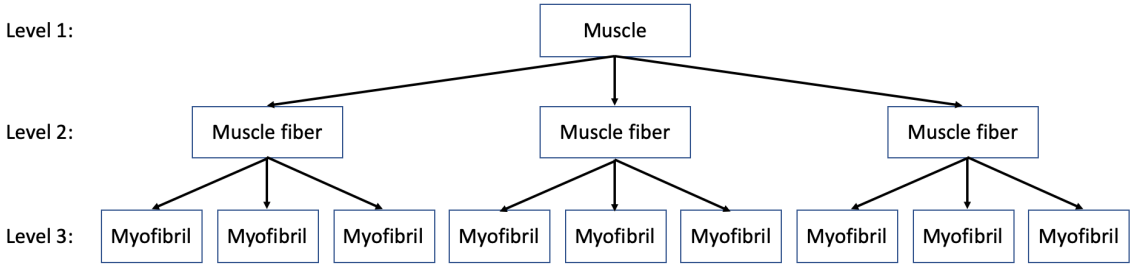

Figure 2: *An example of levels in the human body. Muscles are entities comprising a first level of components (muscle fibers) and a second level of components (myofibrils). Muscle fibers are an entity too, comprising a single level of components (myofibrils).*

The consequences of damage depends very much from the entity at hand. Sometimes, too much physical damage means being dead (e.g., in case of people); sometimes, too much damage means exiting the pool (in the case of a professional having to quit his profession due to reputational damage); sometimes, semi-permanently broken (in the case of a bone); sometimes, permanently broken (in the case of a mechanical component). The concept of damage will become clearer further in the paper. For the moment, the reader can use the following definitions: an entity is said to be macro-damaged if it cannot function as designed, and micro-damaged if some of its components cannot function as designed but the entity itself still can.

When a muscle tries to lift a weight, the load is divided between its fibers and each fiber shares its load between its myofibrils. In general, when a stressor hits an entity, the load is shared between its components at the level below and each component shares its load between its own components, if any.

For the rest of this subsection and of the next one, functional homogeneity and isotropy are assumed (i.e., all components at a given level of an entity are similar and equally connected). Section vi will discuss what happens in heterogenous or anisotropic entities.

When a stressor[4] (a potentially harmful event) hits an entity, the amount of components it damages is linked to its strength. In general, the stronger the stressor, the more components are damaged[5]. Let's see some examples.

- If a stressor hitting an entity is too weak compared to its hardness, it will damage no component. No micro-damage caused.
- If the stressor is just strong enough, it will break the weakest component (even though the components are all similar, say, all muscle fibers, their individual toughness is slightly different and, in nature, normally distributed).
- If the stressor is moderately strong, it will break all components with toughness lower than the individual load supported by each component.[6]

[4] **Stressor**, any event able to cause damage to an entity, if strong enough. Punches, famines, wars and financial crises are all examples of stressors (the first three can cause damage to humans and the last two can cause harm to companies).

[5] This relationship is not linear, though it is weakly monotonic.

[6] Assume a stressor with magnitude L hitting an entity comprising n components at the first level below. Each component will bear a load $l = \frac{L}{n}$. For each component, if l is higher than its individual toughness, that component will break.
This might create a chain reaction. Suppose that the load L is strong enough to break b components. There will be $n' = n - b$ undamaged components left to share the load. Each of these components will now be sharing a load $l' = \frac{L}{n'} > l$. This new individual load might be enough to break a few more components, further increasing the individual load, possibly initiating a chain reaction which will break the entity as a whole. In fact, after a certain b, the increase in individual

- If a stressor is very strong, it will damage a lot of components and the entity as a whole.

The quantitative meaning of "weak", "just strong enough", "moderately strong" and "very strong" will be explored in principle 4. For the moment, the reader should only appreciate the notion that the stronger the stressor, the more components might get damaged, and eventually the entity as a whole.

The attentive reader might have noticed that if an entity already has damaged components, a new stressor hitting it will be perceived as stronger by the remaining components, because fewer of them are left to share the load, each having to bear a stronger individual share of it. This phenomenon will be temporarily ignored for the moment, and addressed further down in the paper, in section v.

load will be higher than the difference in toughness between the last component to break and the next tougher component, assuming individual thoughness is normally distributed.
This process is represented in fig. 3 below.

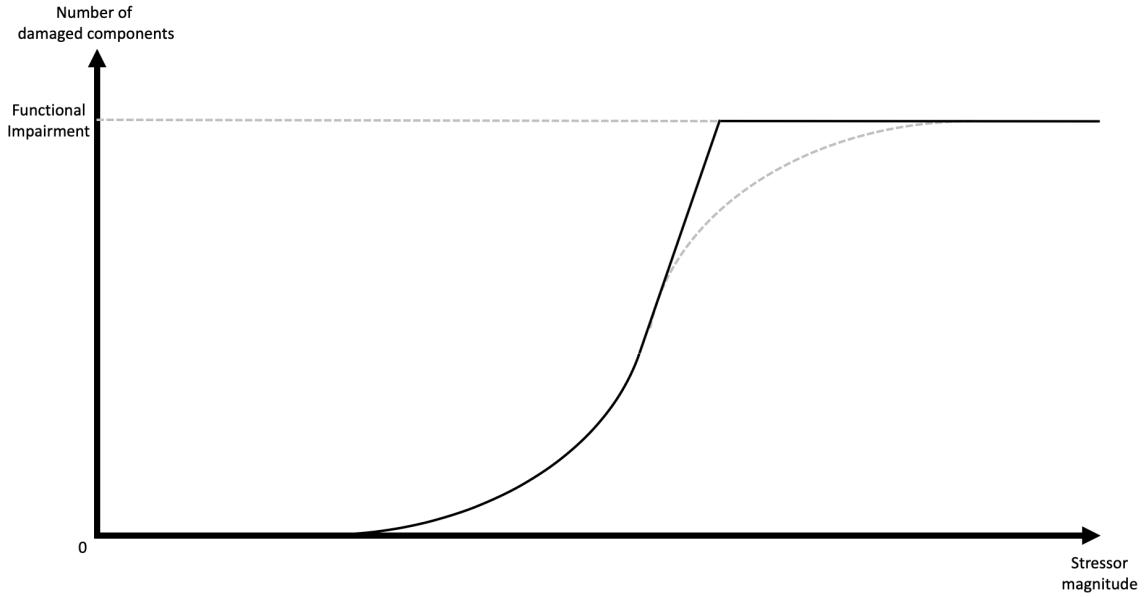

Figure 3: *Number of components damaged at a given level given the magnitude of the stressor hitting it.*

Principle 1. For any given entity, the relationship between stressor magnitude and damage is weakly monotonic.

I.e., all other things equal, stronger stressors cause equal or larger damage.

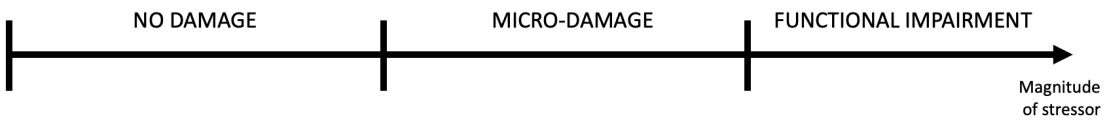

Figure 4: *Damage patterns as a partial function of stressor magnitude*

Functional impairment is defined as damage which prevent an entity from functioning as designed and, in particular, from resisting stressors as designed.

Micro-damage is defined for an entity as damage which causes functional impairment to one or more components of that entity without causing functional impairment to the entity itself.

How much micro-damage causes the entity to suffer functional impairment depends on its brittleness (described in section vi).

"Damage" refers here to any non immediately reversible loss of function to the entity being considered or to any of its components. For example, both a torn muscle and a few torn muscle fibers are forms of damage (the first causes loss of function to the muscle, and the second to the affected muscle fibers only), for both would take at least a few days to recover. Conversely, a temporarily compressed nerve would not be considered damage, as muscle functionality would be regained immediately as soon as the compressing load is removed.

ii. Damage Patterns: the non-ergodicity and scale-dependence of damage

There are simple and complex pathways to damage – ways in which events cause damage. Simple pathways to damage depend on the individual properties of an entity. Complex pathways depend on how entities or their components are connected with each other.

Due to the presence of complex pathways, damage cannot be studied as a purely numerical quantity. When damage occurred, where it occurred and what is it connected to are all information which matter. Hence the need to study damage patterns.

Principle 2. Whether damage causes functional impairment depends on its pattern.

Functional impairment happens if and only if, following damage at some of its components, the upper level is unable to qualitatively function as designed.

How much damage at the components level causes the upper level to stop function as designed depends on the internal structure of the entity being considered, as described in later sections.

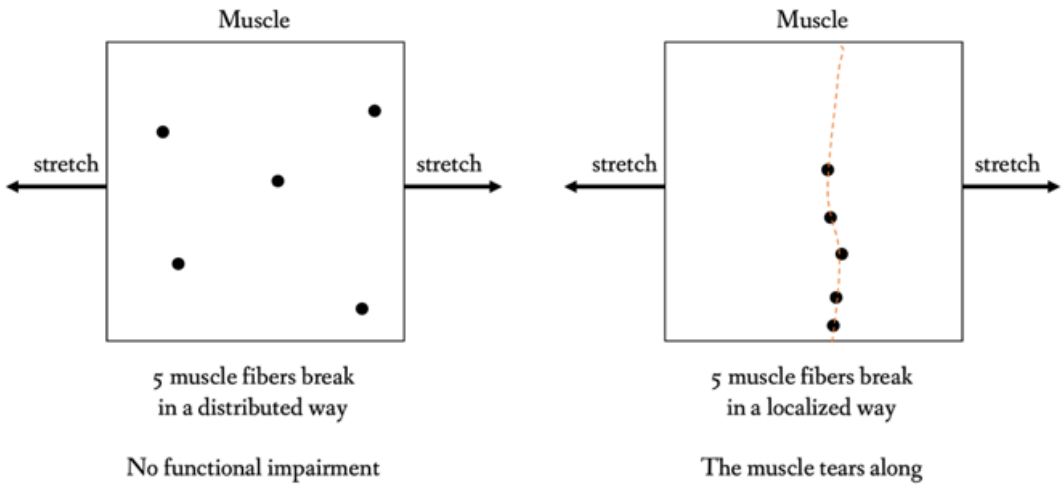

Figure 5: *Damage patterns in muscles (pictorial representation).*

Why is it important? Same quantities of damage might have different consequences. Damage cannot "averaged"; information regarding its pattern[a] has to be retained.

[a]Damage pattern on both the spatial and temporal dimension is important for adaptive entities, whereas the temporal dimension is not important for non-adaptive entities. This point will be justified later in the paper.

iii. Damage and level dependency

Damage[7] is perceived differently depending on the level it is observed from.

Imagine asking a man to lift a load which progressively becomes heavier and heavier. At the beginning, when the load is still very light, there will be no damage to his muscles, to his muscle fibers or to his myofibrils. This is described by point (1) in fig. 6 below. A moderate load might cause damage to a few myofibrils. This would be perceived as functional impairment to the few damaged myofibrils, which are now unable to sustain any load, (2), but would be perceived as micro-damage by the muscle fiber (3), which is still able to carry loads, and as no damage at all by the muscle (4), because no component at the level immediately below is functionally damaged.

Further increasing the load would cause enough myofibrils to break to cause functional impairment to the upper level: the muscle fiber (5). In fact, when enough myofibrils are damaged, the muscle fiber containing them tears. (What "enough" means will be explored in section vi.) This would be observed as micro-damage at the level above, i.e. the muscle (6).

Finally, increasing the load even more would cause enough muscle fibers to break to cause functional impairment to the upper level: the muscle itself (7).

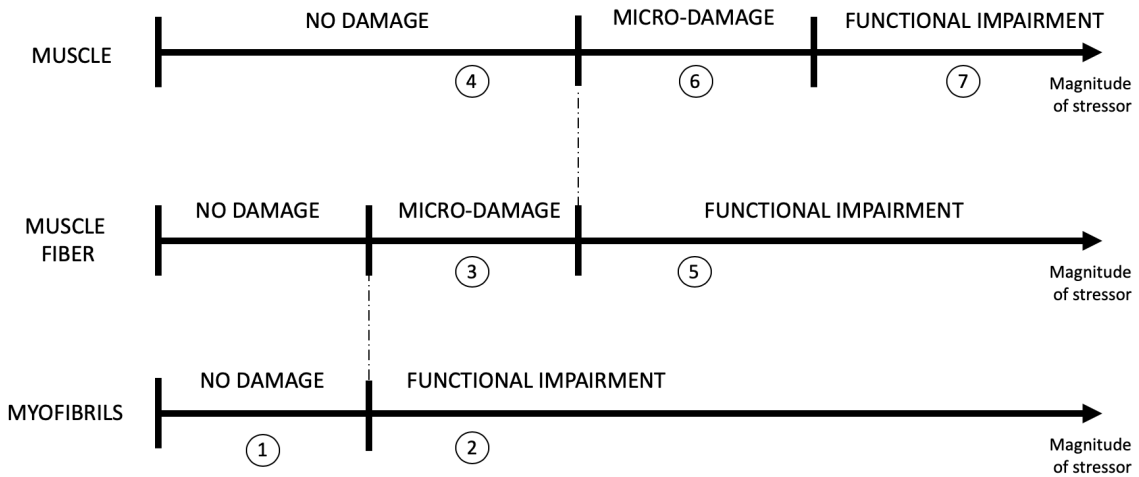

Figure 6: *How damage appears at each level, function of stressor magnitude*

[7] **Damage**, any permanent or semi-permanent modification to an entity or component causing it or any of its components to undergo functional failure. For example, both single muscle fibers torn and full muscle tears are forms of damage from the point of view of the person. Glass cracks and plastic deformations are also forms of damage, whereas elastic deformations aren't – they can immediately be reversed).

Principle 3. Whether damage causes functional impairment depends, between others, on the level considered.

Micro-damage (damage causing no functional impairment) at one level is the expression of damage causing functional impairment at the levels below.

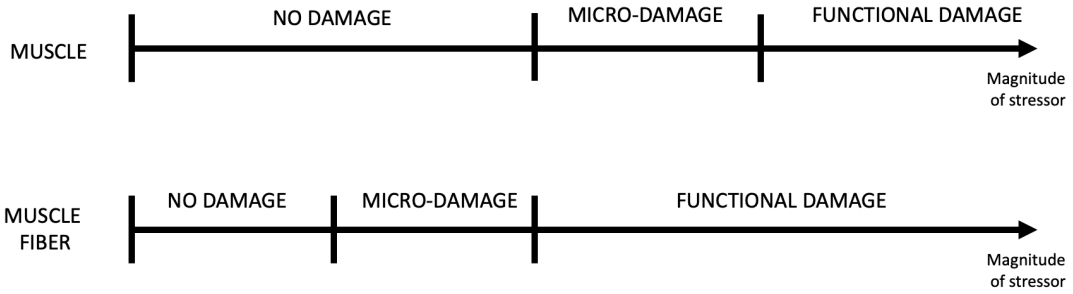

Figure 7: *How damage appears at each level, function of stressor magnitude*

Why is it important? As will be explained in the next section, for adaptive entities damage is both desirable and undesirable. In particular, damage at lower levels is desirable, in a measure. Treating all damage equally regardless of the level it takes place at is a dangerous simplification.

iv. Thresholds: how much is "a strong hit?"

It is possible to define two thresholds which use stressor magnitude to discriminate between damage patterns, as shown in fig. 8.

The first threshold is called *First Damage Threshold* and corresponds to the magnitude of the weakest stressor able to cause micro-damage: functional impairment at one of the components of the entity being hit. A few muscle fibers torn – the normal consequence of a gym session – is an example of micro-damage.

The second threshold is called *Functional Impairment Threshold* and corresponds to the magnitude of the weakest stressor able to cause functional impairment: damage causing the entity being hit to stop functioning as it should. A tear at a muscle following having tried to lift a too-high load is an example of functional impairment: the entity as a whole – the athlete – is unable to function normally for a few days.

The two thresholds have to be measured empirically, for both depend on complex interactions between the components of the entity.[8] [9] [10]

Principle 4. The stressor magnitude thresholds separating damage patterns are empirically determined.

The *First Damage Threshold* is defined as the magnitude of the weakest stressor able to cause one of the components of the entity being hit to become functionally impaired.

The *Functional Impairment Threshold* is defined as the magnitude of the weakest stressor able to cause functional impairment to the entity being hit.

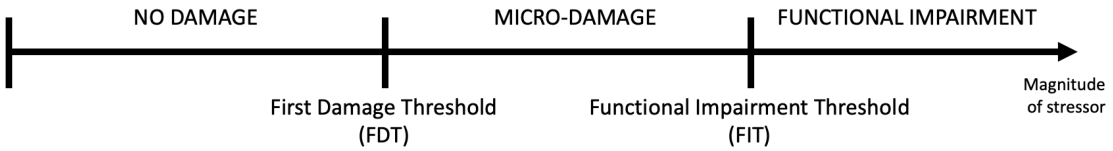

Figure 8: *The First Damage and the Functional Impairment Thresholds*

Why is it important? The position of the thresholds change following damage. Believing that they are the cause rather than the consequence of reactions to damage leads to wrongly predicting future behavior after an entity gets damaged (and consequently changes its internal properties).

In general, as shown in fig. 9, the First Damage Threshold at a given level corresponds to the Functional Impairment Threshold at the level immediately below. This is self-evident from the

[8] Whereas it is possible to use models to deduce the thresholds of extremely simple entities organized with extreme regularity (e.g. mineral crystals found in nature), it is impossible to do so for most entities encountered in the real world.

[9] For example, given an entity and a stressor, whether the former will be damaged by the latter depends on the internal structure of the entity, eventual internal redundancies and their locations, the shape, etc.; only considering the structure of a single component would not be enough.

[10] The thresholds are usually fuzzy, as different factors influence the magnitude of the stressor (for example, the angle with which the hammer hits the rock).

definition of the First Damage Threshold: the magnitude of the weakest stressor able to cause functional impairment at one of the components of the entity being hit.

Principle 5. The First Damage Threshold at any given level corresponds to the Functional Impairment Threshold at the level below.

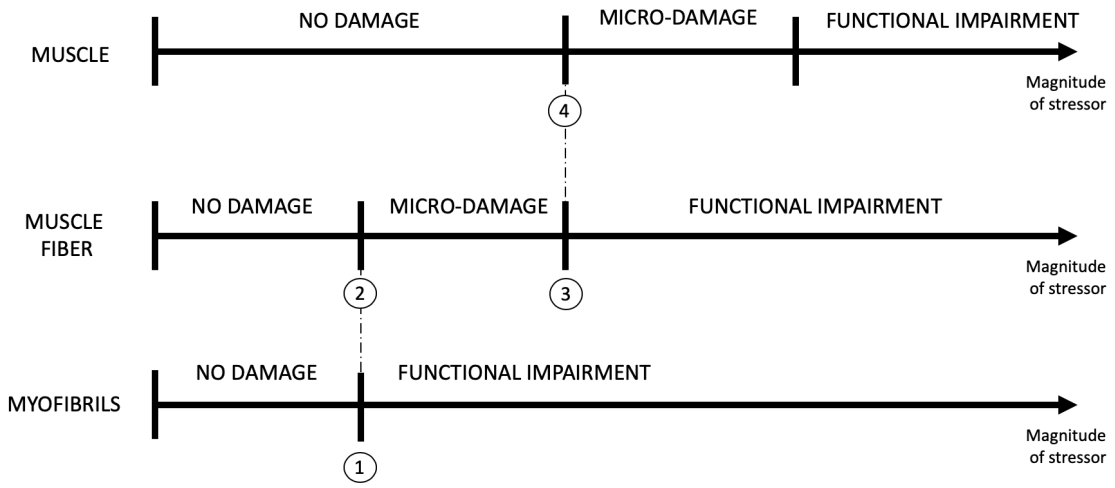

Figure 9: *The Functional Impairment Threshold at one level corresponds to the First Damage Threshold at the level above.*

Why is it important? This concept will be crucial later on to understand how the movement of thresholds at different levels are correlated.

Let's examine the cause-effect relationship between thresholds and damage patterns. As previously mentioned, thresholds are an abstract concept used here to provide clarity in how entities react to stress. There is no single absolute threshold that determines the kind of damage an entity will display. The real process is bottom-up and works as follows:

1. A stressor hits the entity.
2. If the entity comprises components (cells, crystals, etc.), the stressor gets distributed amongst them, level by level, with each component absorbing part of it depending on the internal structure.
3. Each component, based on its own mechanics, might get damaged or not (muscles fibers might tear, grains might separate in ceramics, and so on, depending on the class of material at hand).
4. If no component broke, then it means that the stressor magnitude was below the First Damage Threshold for the component.
5. If some component broke in a distributed pattern without functional impairment at the upper level, then it means that the stressor magnitude was above the First Damage Threshold but below the Functional Impairment Threshold for the component.

6. If many components broke in a localized pattern causing functional impairment at the upper level, then it means that the stressor magnitude was above the Functional Impairment Threshold for the component.

In other words, it is the bottom-up reactions of the entity that determine the thresholds; not the other way around.

v. Fatigue: accumulated damage

In materials science, fatigue is the accumulation of micro-damage until the point the entity breaks, even if it had never been subject to stressors above its Functional Impairment Threshold. For example, hitting a rock with a hammer 100 times, each time with a hit of moderate strength (comprised between the two thresholds), would cause micro-cracks to accumulate and to expand, until with the last hit the rock breaks open.

When an entity breaks due to the accumulation of micro-damage, it is said it broke *due to fatigue*. In the rest of the paper, the term *fatigue* will be used with this meaning, and never with the common meaning of "being tired".

1000 hits with magnitude below the First Damage Threshold of the rock entity being hit would do absolutely nothing towards fatigue. Only hits from a stressor above that thresholds would count against fatigue.

Accumulated fatigue decreases the First Damage Threshold and the Functional Impairment Threshold of the components at the level at which fatigue is accumulated and at every level above.[11]

In non-living entities, fatigue accumulates because there is no process taking care of repairing or substituting the broken components. In living entities, there are such processes ("healing"). However, living entities can also accumulate fatigue, if the rate at which micro-damage is caused is higher than the rate at which it is healed. For example, a man running on tarmac once a month might cause damage to his knee joints but not suffer any long-term consequence, if the healing rate is such that by the next time he runs, micro-damage from the previous session has been healed. Conversely, a man running once a day is likely to cause accumulating damage to his knee joints, as healing cannot keep up with the damage accumulation.

[11] A stressor with magnitude L hitting an entity made of n components is shared between them and perceived by each as a load $l = \frac{L}{n}$. If fatigue has been accumulated as b damaged components, only $n' = n - b$ undamaged components are left to share the load, and each of them will be stressed by a load $l' = \frac{L}{n-b} > l$. Because L is now perceived as l' instead of l, a lower value of L will be required to damage the weakest component. That value will be the new First Damage Threshold, lower than the original one.
Similarly, the Functional Damage Threshold will decrease too.

Principle 6. Micro-damage accumulates if generated at a faster rate than the healing rate of the entity (if any), increasing the likelihood that future stressors damage the entity.

Any hit above the Fatigue Threshold causes micro-damage which counts towards fatigue.

The Fatigue Threshold of an entity corresponds to the First Damage Threshold of its lowest level of components.

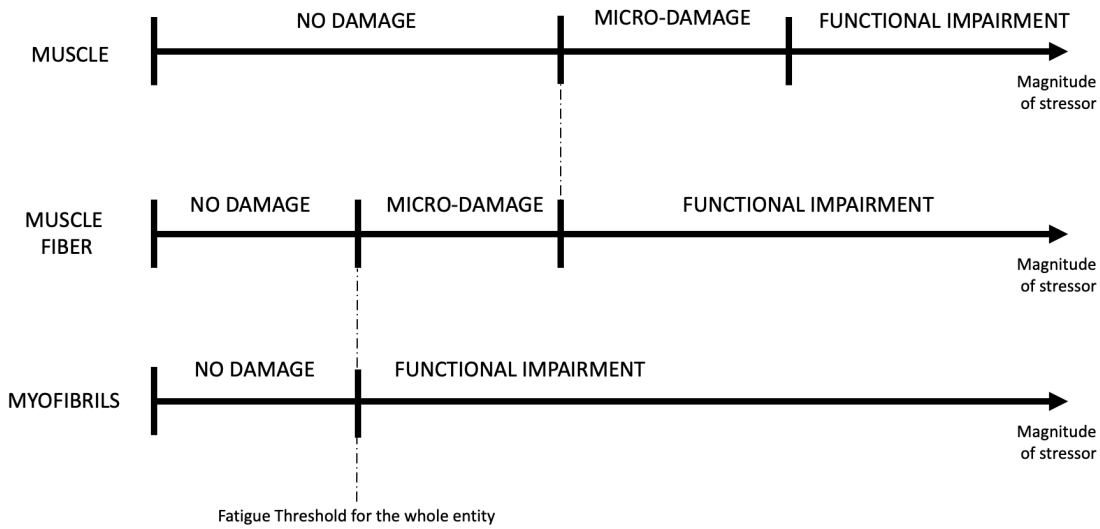

Figure 10: *The Fatigue Threshold*

Why is it important? There is a limit to how much micro-damage is good, even for entities able to heal. Too often, athletes understand this only after a painful chronic injury.

vi. Brittleness: space- and time-concentration of damage

Imagine a man hitting a glass-made window with his hammer, with varying degrees of strength. The man would only be able to either hit the glass so lightly not to cause any damage, or to hit it strongly enough to break it. The glass is so brittle that it is almost impossible to cause damage to it without also breaking the glass.[12]

Brittle entities – such as those made with glass or ceramic – are characterized by the first micro-crack expanding almost immediately into a full-crack breaking the entity. The cause of brittleness will be described in the next subsection.

Brittleness can both be considered a binary concept or a non-binary one. In the first case, an entity is brittle if and only if any damage no matter how little would cause the entity to break.

[12]Modern windows and windshields *can* show a crack without breaking. This is usually due to the glass being covered on at least one side with a layer of a polymer which "keeps the glass together" in case of small cracks. The paper uses the term glass to refer to the basic material used during most of recent human history.

This means that the First Damage Threshold and the Functional Impairment Threshold coincide. In the second case, it can be said that "an entity is brittler than another" if the distance between the two thresholds is shorter in the former than in the latter.

Generally, in brittle entities, the probability that the second point of damage is adjacent to the first one is very high. For example, in glass – a brittle material -, it is extremely likely that the second separation between glass grain borders will be adjacent to the first one. Instead of creating a second micro-crack, the first one will be expanded.
The brittler the material, the more likely that the second point of damage will be adjacent to the first one.

Entities can be hard or soft regardless of whether they are brittle. In other words, their First Damage Threshold might be high (as in diamond) or low (as in glass). Whether an entity is brittle only depends from the relative distance between the two thresholds – it doesn't depend on their absolute position.[13]

Principle 7. The brittler the entity, the closer the Functional Impairment Threshold is to the first damage threshold.

Often, brittleness could be estimated using as a proxy the relative probability that the second component being damaged in an entity is functionally adjacent to the first one, compared to the probability of it being a random component of the entity (where "functionally adjacent" means that it contributes to sharing the additional load provoked by the first component failing more than an average component).

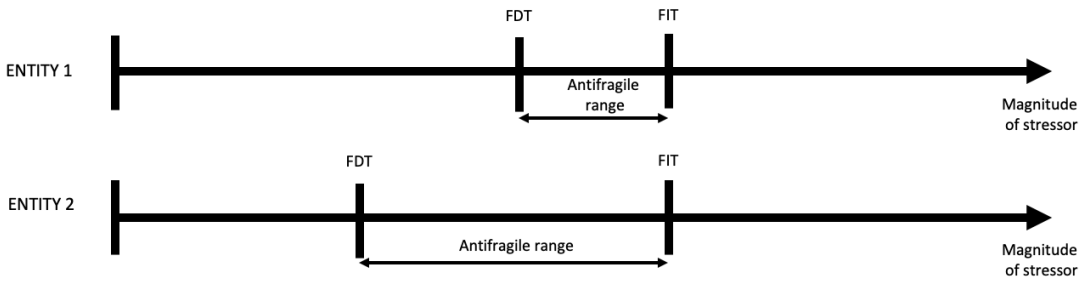

Figure 11: *Here, the first entity is brittler than the second, as its thresholds are closer to each other and thus its antifragile range is shorter. Note that both entities have the same Functional Impairment Threshold: brittleness is about "little damage causing failure", not about "how strong a stressor can be resisted before failure".*

Why is it important? Understanding how brittleness brings the FIT closer to the FDT will be crucial for visualizing how thresholds change in response to changes in internal structure.

It might seem that given two entities with the same toughness (the position of the Functional Impairment Threshold), the brittle one is less likely to break, because its First Damage Threshold is higher, reducing micro-damage and thus decreasing the likelihood to break due to fatigue. In

[13] Readers versed in materials science might have noticed a relationship between the First Damage Threshold and hardness, the Functional Impairment Threshold and toughness, and between the distance between the two thresholds and resilience.

practice, the opposite happens. Entities (if alive) or people using entities (if the entities are objects) use micro-damage to estimate vulnerability and adjust risk-taking accordingly. Given the two entities described at the beginning of this paragraph, the less brittle entity is more likely to suffer micro-damage, which would signal vulnerability, causing it or the people losing it to reduce its exposure to the environment by being more cautious, reducing the stressors it will be subject to in the future. This will be addressed more in detail in section Exposure Dynamics: Why People Take Risks.

Principle 8. Brittleness prevents clues of vulnerability from appearing before failure.

Often, the position of the First Damage Threshold is used to estimate the position of the Functional Impairment Threshold. This leads to overestimating the Functional Impairment Threshold in entities whose brittleness the observer is unaware of.

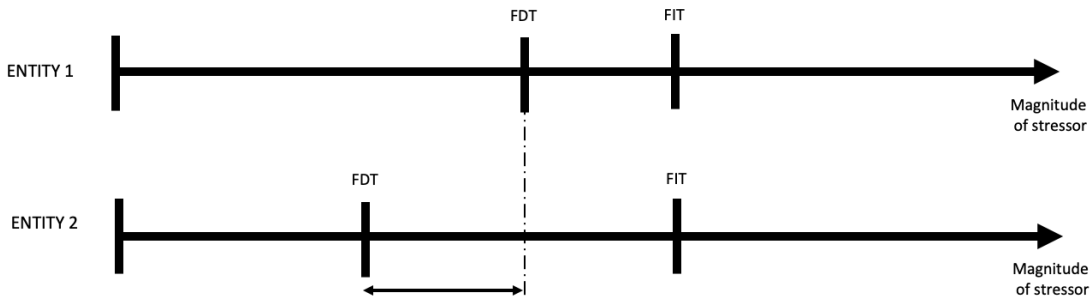

Figure 12: *In the range highlighted by the double arrow above, stressors will cause observable damage only in Entity 2. Entity 1 is therefore likely to be judged as having a higher Functional Impairment Threshold by an observer who only tests stressors lower than the Functional Impairment Threshold.*

Why is it important? Given two non-adaptive entities with the same Functional Impairment Threshold but with the first having higher First Damage Threshold than the second, in theory the first would survive the environment longer (because fewer stressors would be causing micro-damage which accumulates); in practice, the second one survives longer, as the micro-damage it receives signals the necessity to handle it with more care.

vi.1 Brittleness and redundancy

At a first analysis it might seem that the cause of brittleness is that a stressor strong enough to cause first damage is also strong enough to cause functional impairment. While empirically true, the previous statement is the result of brittleness, not its cause. **Entities are brittle if the components near the ones which are broke cannot adapt to share some of the load which would have been carried by the broken one.** In materials which are not brittle, such as metals, the regions around an initiated crack deform to contribute to load sharing.[14] In entities other than materials, neighbors use some form of internal redundancy to adapt and help share the load. For example, a company in which workers are willing to make some extra effort to take over the workload of a colleague

[14] In the case of non-brittle materials, deformation helps because it makes the end of the crack round rather than sharp. Sharp cracks need very little force to propagate – as anyone sharply folding a paper before splitting it in two knows.

who fell ill is less brittle than companies in which one absent worker (the "first damage") would cause some of the company's processes to stop ("the functional impairment"). In brittle entities, bottlenecks and failure points have nonlinear negative consequences.

Elastic and plastic adaptation Redundancy can be considered capacity for elastic adaptation: a temporary form of adaptation to stressors. A supermarket that has a stocked warehouse (the redundancy) is able to accommodate temporary delays from the supply trucks.

The next section will describe antifragility: capacity for plastic adaptation: a semi-permanent form of adaptation to stressors (only partially permanent, until the next few steps of adaptation, but it doesn't revert immediately as the stressor is not applied anymore, as elastic adaptation does).

The level at which redundancy is stored matters . Redundancy affects differently the brittleness of components at different levels, depending on their position relative to redundancy.
Consider the brittleness to famine that a city might have. Imagine a town with a single supermarket stocked with food (redundancy at the town level) whose habitants do not stock food at home, because they can source it easily from the supermarket on a daily basis. In case, say, a snowfall prevents the supply truck to bring food to the supermarket, none of the households will notice that the truck didn't arrive until they buy and finish the stock already in the supermarket. Thanks to the supermarket stock, the households are resistant to moderate delays in the supply truck delivery. But, if the snowfall continues to the point that one household suffers famine, it is highly probable that very soon a second household will also suffer famine, for the central supply of redundancy (the supermarket) is emptied at the same time for all households (they all had no stock so they all relied on the same source of redundancy – the supermarket). The stocked supermarket does makes the town more resistant, but not any less brittle. It increases the First Damage Threshold, but it does not affect the distance between the First Damage Threshold and the Functional Impairment threshold.[15]

Imagine instead a town with a tiny supermarket which stocks almost no food; the households stock food at their houses instead. In case of a snowfall, the first household to suffer famine – the one with less stock or more mouths to feed – will do so much earlier. However, the time between the first and the second household suffering famine will be much larger than in the first example, as different households tend to have different levels of stock. The town with the redundancy at a lower level is bound to suffer more micro-damage but also is less brittle (less chances of becoming fully wiped out by a famine, *given one household suffering from it*).

Which town is more resistant? Under a static analysis, the first one, as it accumulates less fatigue damage. Under a dynamic one – which will be the object of the next section – the second town, as suffering a few deaths from famine (these coming from a stressor with magnitude just above the First Damage Threshold) will make it reconsider its supply chain and stocking choices.

Redundancy at a given level only makes that layer and the levels above less brittle, but it does not influence the brittleness of eventual levels below. It does increase the First Damage Threshold of every level, though, as stressors stressing a given level including a redundancy would partially be absorbed by the redundancy before being passed down to the components at the level below.

There are multiple causes for an eventual lack of redundancy at the component level; the most common ones are usually rigidity, tension, specialization and centralization.

[15] Actually, once second-order effects are taken into the account, the well-stocked supermarket does affect the Functional Impairment Threshold: it decreases it. This is because the more the supermarket is *apparently* reliable, the less stock will the households hold, making the town brittler to supply chain breakdown.

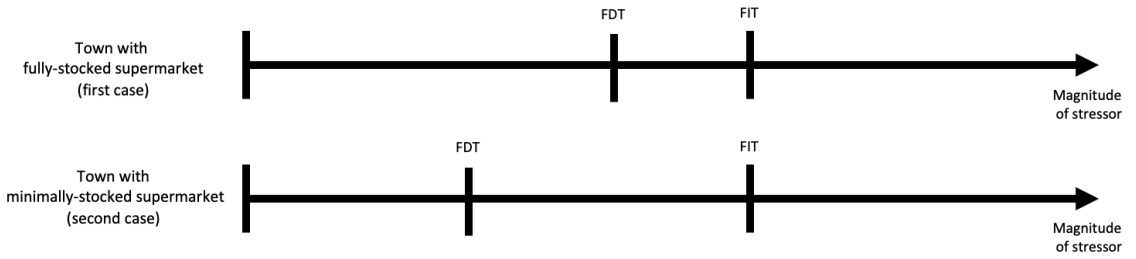

Figure 13: *A chart describing the example above.*

Rigidity: some components are naturally and inherently rigid, unable to change their structure or functioning to accommodate workload different from the specific one they are usually responsible for. For example, the grains forming ceramics are almost inelastic and thus unable to bend in order to share the load coming from a nearby crack. As another example, the employees of an office might be very rigid and unwilling to do anything to accommodate the void left by an employee who is sick; as a result, the workload which would have been processed by the sick employee is not absorbed by its colleagues and "overflows", causing the office to not function anymore as desired or, in the worst cases, risking to damage the office or the other employees (an ineffective office might, for example, become the target of downsizing after customers flock to more reliable competitors).

Tension: it might happen that components usually able to adapt and accommodate extra work are unable to do so if under tension. For example, the office workers of the previous example might be willing in normal conditions to do one hour overtime to cover their sick colleague, but might be unwilling to do so if they are already usually working overtime due to chronic inefficiency or due to an ineffective manager who constantly overworks them in the name of "efficiency".
An interpretation for this cause of brittleness is that the pre-existing tension uses up some of the available redundancy, and little is left to accommodate extra stress.

Specialization: in an homogeneous entity, all components have a similar capacity to share load. Every component within a reasonable distance from one or more which are damaged contributes to sharing the additional load arising from the damaged components not being able to perform any work anymore. Conversely, in an heterogenous entity, components are specialized and differ in their ability to absorb work. For example, in a company, if an accountant is sick, only other accountants can contribute sharing the workload. Consequently, two accountants being sick at the same time result in bigger problems than, say, one accountant and one customer service rep being sick at the same time.

If components are specialized, when one of them breaks, only the redundancies of these with the same specialization can be used to absorb the increased tension. Load sharing is therefore lower and slower, making the entity brittler.

That's not to say that specialization cannot make the entity stronger: it is often the case. But it always makes it brittler too.

Centralization: in section vi, it has been said that brittleness could be expressed as the probability that the second point of failure is adjacent to the first one. Specialized entities tend to group together components with similar functions (often "efficiency purposes", forgetting that with $t \to \infty$, efficiency $\to$ survival) and become centralized.

Centralized entities are especially vulnerable to geographically-concentrated stressors. For example, if all accountants of a company work in the same room, the probabilities that given an accountant catching a flu, the second employee falling sick is an accountant too are extremely high. If that happens, the remaining accountants will struggle to cover the workload of two missing people – no matter how many employees of another kind (say, customer service reps) are in the company. The more centralized an entity is, the brittler it is.

Principle 9. Brittleness is caused by lack of redundancy. Redundancy at a given level makes that level and the ones above less brittle (but makes the ones below brittler, due to fragilization, as explained in "Adaptation Mechanics").

The main causes for lack of redundancy are:

- Rigidity (lack of inherent redundancy)
- Tension (lack of unused redundancy)
- Specialization (lack of functional redundancy)
- Centralization (lack of geographical redundancy)

Why is it important? Too often, redundancy is considered a waste. It often is not, just as no one would consider having bought a fire extinguisher a waste if he didn't get to use it. Redundancy is useful for reducing brittleness.

vi.2 Tension-caused brittleness

In [Taleb, 2012], Nassim Nicholas Taleb notices the following. In metallurgy, annealing is the process by which a metal piece gets exposed to high temperatures for a long time, so that the heat breaks the pre-existing bounds between atoms and allows them to reconfigure into a new internal structure with less internal tension and thus less hard and less brittle. The same happens in populations of humans: the freer they are to reconfigure, the more stable and the less brittle the resulting configurations.

vi.3 Brittleness as damage concentration over space and time

In this subsection so far, brittleness has been presented as the effect of internal structure on the concentration of damage evaluated on the space dimension. However, brittleness also causes concentration of damage over the time dimension. Due to their likelihood to either suffer no damage at all or to break at once, brittle entities usually experience a (relative) lifetime of no damage followed by an almost immediate transition to being broken, without suffering the periodic micro-damage events that less brittle entities suffer.

Principle 7 presented the notion that brittleness could be estimated using as a proxy the relative probability that the second component being damaged in an entity is functionally adjacent to the first one, compared to the probability of it being a random component of the entity. In other

words, that brittleness is linked to the spatial concentration of damage. Similarly, brittleness could be estimated using as a proxy the relative probability that in the instant following a component being damaged another component gets damaged, compared to the probability of a component being damaged at any point of time. In other words, brittleness is linked to the time-concentration of damage.

Another formulation would be that brittleness makes damage concentrate in space and time.

II. Adaptation Mechanics: How Entities Adapt

i. (Anti)fragility

Nassim Nicholas Taleb first coined the term "Antifragile" to refer to entities that increasingly benefit from volatility [Taleb, 2012] (and was definitely the main influence for this paper, including the following examples). For example, human muscles are antifragile, because lifting 1kg 100 times causes almost no benefit nor harm, but lifting 10kg for 10 times causes strengthening (muscle growth). Conversely, most artificial entities exhibit a negative non-linear response to all stressors[16] (they are harmed by volatility). For example, a teacup does not break if it falls 100 times from 1cm height, but it breaks if it falls once from a 100cm height: it is a fragile entity.

Principle 10. A definition of (anti)fragility

An entity is *antifragile* to volatility if it increasingly benefits from it.

An entity is *fragile* to volatility if is is increasingly harmed by it.

Note: in this manuscript, the term "strength" is used to indicate the capacity to resist stressors. Readers versed in material sciences might wonder whether it refers to the specific terms of hardness, resilience, or toughness. For reasons that will become apparent later in this section, in antifragile entities it refers to the three of them, as they are deeply intertwined and always change together, unless the internal structure of the entity is changed too.

ii. The requirements for antifragility

Antifragility is the response to a stressor causing some form of partial damage, triggering some form of adaptation. The details of this will be seen later. The only way an entity can be partially damaged is to comprise multiple components and have only some of them damaged. For example, the antifragile reaction of muscle growth follows the tear of just a few muscle fibers while lifting weights. A monolithic entity cannot be antifragile, because the monolithic is either intact or broken; it can only be fragile.

Moreover, an antifragile entity must possess some form of re-growth. Otherwise, the percentage of its components which are broken would progressively and monotonically grow until the entity as a whole would break (as described in subsection "Fatigue: accumulated damage"). Some form of re-growth of new components has to be present, though it doesn't have to be provided by the entity itself (it could be parasitic or symbiotic). For example, it could be argued that the Internet is

[16] At least, to those above their hardness (their First Damage Threshold).

an adaptive entity, even if the capacity to grow servers and websites comes from developers, an external entity.

This re-growth might be constant (such as human populations, where children are constantly conceived) or triggered/augmented by harm (such as muscle fibers, which grow in number when some nearby are teared). [17]

Principle 11. An entity is antifragile if and only if:

- It comprises multiple components or sub-entities
- The components can break independently
- New components are generated continuously or at least in response to other components being damaged.

ii.1 Damage as the signal directing adaptation

The most common example used for antifragility through this paper is muscles: lifting weights damages muscle fibers and triggers their regrowth, in overcompensation, so that the next time we will be able to lift heavier weights. However, it is not certain that breaking muscle fibers directly triggers growth – perhaps other pathways are at play, such as the brain inferring the need for strength and triggering hormones. We don't know yet with certainty.

What we do know, is that antifragile entities adapt to what damages them – even thought damage is not the direct trigger. For example, muscles grow in response to actions which cause them to get damaged – regardless of whether damage is the trigger. This makes sense, because what has the potential to damage an entity usually also has the potential to threaten its survival, so the entity should adapt to it.

Hence the importance of studying damage as the signal which directs adaptation.

iii. Fragility

An entity is said to exhibit a fragile response to stressors if damage from the stressor grows exponentially with magnitude.

Fragile responses are caused by the inability to have some components damaged without this causing others to be damaged as well. This is usually the result of one of the following two root causes: being monolithic or having an internal structure prone to systemic risk. An example of the latter case would be a city where all the warehouses are situated near the river. While the warehouses are apparently independent, a single flood causing damage to a small percentage of the city might wipe out all warehouses, exposing the whole population to famine.

No antifragile entity exclusively reacts to stressors in an antifragile way. Even antifragile entities can exhibit fragile responses.[18] For example, humans are antifragile and lifting weights

[17] In general entities are at competition with each other in inverse measure to how much "DNA" (genetic or cultural) they share. The more they are at competition, the more it makes sense to attempt continuous growth rather than saving resources. For example, muscle fibers share DNA so they only grow when triggered by stressors.
This observation is not a rule, though, as exceptions exist.

[18] Page 271 of [Taleb, 2012].

causes them to become stronger, but attempting to lift a too-heavy weight would cause a fragile response (for example, a muscle strain or an hernia). Therefore, **antifragility is always related to both an entity and to a range of stressors.** Antifragile entities exhibit no response to stressors below the First Damage Threshold[19], they exhibit a fragile response to stressors above the Functional Impairment Threshold[20], and they exhibit an antifragile response to stressors between the two magnitude thresholds, as illustrated in fig. 20 below.

Principle 12. Every antifragile entity can exhibit fragile behavior if hit by a strong enough stressor.

The antifragile behavior is expressed only for stressors whose magnitude is between the First Damage Threshold and the Functional Impairment Threshold.

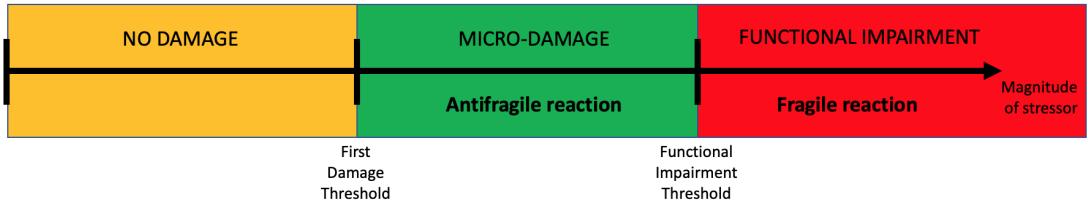

Figure 14: *How antifragile entities reacts to being hit by a stressor.*

Fragile entities react similarly, but do not have a range of stressors producing an antifragile response.

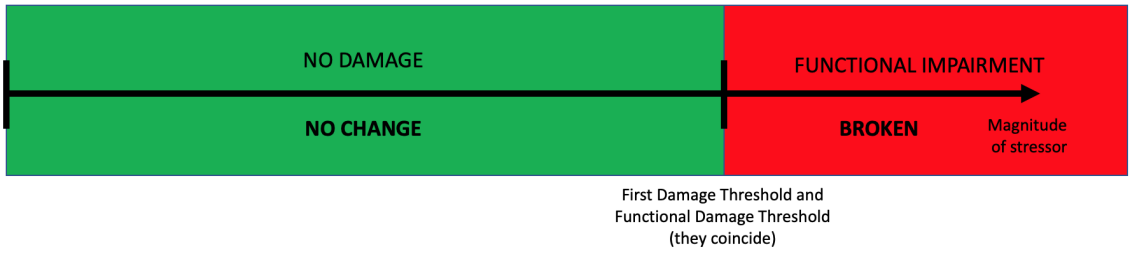

Figure 15: *How fragile entities reacts to being hit by a stressor.*

iv. Brittleness and fragility

An entity is fragile if any of the following are true:

- It is monolithic
- It is brittle (i.e., damage to its components is immediately followed by functional impairment at the entity level)

[19] **First Damage Threshold**, the magnitude of the weakest stressor able to cause one of the components of the entity being hit to become functionally impaired.

[20] **Functional Impairment Threshold**, the magnitude of the weakest stressor able to cause functional impairment to the entity being hit.

- It lacks the ability to (re)grow its components.

Therefore, brittleness implies fragility, but fragility does not imply brittleness.For example, a piece of cheese is fragile but not brittle: it can suffer many cuts before becoming unable to sustain any load (thus it is not brittle) but cuts accumulate without repairing (thus it is fragile).

v. Two patterns of growth: organs and organisms

Note: this subsection is an inexact simplification which covers most but not all cases present in nature. It is presented here to give the reader an incomplete overview of some of the patterns through which antifragile entities grow their components and it should not be considered precise nor comprehensive.

Antifragile entities grow their components following one or both the following patterns.

Adult **organisms** constantly attempt to grow their numbers in excess of the replacement rate. For example, a human population constantly attempts to have children (of course, only adults within a given age range do attempt so; however, most adult organisms feel the impulse to procreate with external conditions exerting a limited influence on it) and usually tries to do so above replacement rate (only in recent years and only in richer countries, the number of newborns per couple fell below 2).

Adult **organs** only grow when needed. For example, muscles only grow new myofibrils when exercised, and atrophy if not used for long time.

Organisms part of the same entity (for example, a tribe or a species) do not have the same DNA and therefore compete with their peers. It is in their individual DNA's interest to continuously grow offspring, even when resources are scarce.

Organs part of the same entity (for example, muscles and bones belonging to the same person) share the same DNA. Therefore, they have no interest in growing their numbers in conditions in which doing so would decrease the survival probabilities of the host – for example, growing too large muscles helps in case of war but might be detrimental in case of a famine.

It has to be noted that few real world entities fall completely in one of the two categories above. In many cases, the following is often representative of observed behavior: entities adapt like organisms in the measure they do not share DNA (genetic or cultural) and adapt like organs in the measure they do. However, the content of this subsection is not a rule and has exceptions; it has been presented just as a consideration to help the reader entertain the notion that there are two complementary processes underlying (re)growth. Readers failing to internalize the above often believe that adaptation is synonymous with evolution – it is not; it is only one of the ways entities adapt.

In organs[21], adaptation happens when they atrophy in correspondence to a need to atrophy and when they grow in correspondence to a need to grow. For example, losing muscle during long winters and growing muscles during hunting seasons are both adaptive behaviors. So are companies growing headcount while the economy grows and cutting personnel while the economy shrinks. Conversely, maladaptation happens when organs atrophy while the need to be strong is there or when they grow more than the environment could sustain them. For example, a modern adult whose muscles shrink due to a sedentary lifestyle maladapts to his environment, for he becomes weak in case of a rare event requiring muscles (such as carrying a heavy suitcase over

[21] or, to be more precise, in entities that behave like an organ, in the measure in which the behave like an organ.

a flight of stairs)[22]. Similarly, a car manufacturer which grows its factories too much due to a sudden increase in cars sold might find itself in a bad position if the trend stops.

In organisms, adaptation happens when stressors from the environment damage the unfit ones, causing the fit ones to reproduce faster relatively to the unfit. Maladaptation happens when a stressor proper of the environment does not manifest for long enough[23], causing the population to adapt to the absence of that stressor. For example, a few short winters might be enough to "punish" the households which stocked too much food and to "reward" these which stocked less food or ate more of it. This might cause the population to adapt to short winters, stocking less food and becoming more vulnerable to an eventual long winter. Adapting to short winters coincides with maladapting to long ones. Whether it is an adaptation or a maladaptation depends on whether the trend of shorter winters was representative of a changing environment or the result of a spurious coincidence.

Maladaptation is the response to exposure to prolonged a non-representative sample of the environment.

"Prolonged" in the sentence above means for a period of time comparable to the period it would take for the entity to considerably change the phenotypes (features) expressed by its "DNA" (genetic, cultural or technological). A genetic DNA that can change rapidly (due to a high rate of mutations or to a short reproductive cycle) is prone to both adaptation and maladaptation. Similarly, a cultural DNA that can change rapidly (due to a lack of traditions or of respect for them) make a population prone to both adaptation and maladaptation.

The adaptive mistakes absence of evidence (of harm) for evidence of absence because false negatives are just as expensive as false positives in conditions of scarcity of resources, where missing an opportunity to obtain some of them would mean not having enough to mitigate other risks. This topic will be further explored in section IV.

vi. (Anti)fragilization: convexity changes over time

The chart of principle 12 is not static. As an antifragile entity undergoes micro-damage, its capacity to withstand future stressors increases. This process is called antifragilization and is described below. An example would be how, after a gym session having caused muscles to grow, the athlete is able to benefit from lifting weights which a few weeks before would have been too heavy and would have injured him.

When an antifragile entity suffers micro-damage (damage at its components which does not cause functional impairment at the entity level), one or both of the following take place:

- More components are produced as a form of overcompensation (as a way to build a redundancy of strength necessary to withstand future stressors) [Taleb, 2012].
- The weakest components are damaged and won't reproduce: the "gene pool" gets better.

In the case of entities in which the capacity for growth is externally provided (for example the internet, in which servers and websites are designed by tech companies and developers), micro-damage causes one or more of the following effects:

- The weakest external providers of growth exit the pool (e.g., the incompetent software developer who designed the broken website is fired) and better ones are hired.

[22] Lack of muscles is probably also bad for its metabolic effects, such as their effects on insulin levels.
[23] How much is "long enough" will be clarified in section III.

- The weakest design patterns get abandoned (e.g., the server designers stop using a given component) and replaced with better ones.
- A mix of the previous two points: the external providers abandon their weakest patterns (e.g. a software developer stops designing websites in a given way) and adopt new, better ones.

All of the five processes listed above generate strengthening: after an adaptation period, the entity becomes stronger and less brittle.

Principle 13. Micro-damage causes strengthening in antifragile entities.

An antifragile entity undergoing micro-damage, after an adaptation period, becomes able to withstand stronger stressors.

Why is it important? Micro-damage has a required intake for antifragile entities to grow strong and stable, just like nutrients and vitamins.

Similarly, lack of micro-damage in an antifragile entity causes weakening[24], following one or more of the processes below:

- The weakest components survive even if they are unfit and get to reproduce – natural selection isn't working properly anymore.
- The lack of damage from a given risk is interpreted as a lower need to use precious resources to build redundancy against that risk, resulting in those resources used elsewhere or with another purpose.

The results of lack of micro-damage go under the names of maladaptation, decadence, and recklessness: all result in a weaker entity, less able to withstand future stressors.

Principle 14. Lack of micro-damage causes weakening in antifragile entities

For the antifragile, "Use it or lose it", [Taleb, 2012].

Why is it important? The modern urge to guarantee comfort at all times makes us weak. It's not comfort which is the problem, but absence of discomfort.

In principle 11, it has been written that an entity is antifragile if and only if it is made of components which can individually be damaged. This affirmation can be proven by observing the behavior of what is not made of components: the monolithic.

A monolithic entity cannot suffer micro-damage – damage at *some* of its components – because it contains a single component: itself. It can only maladapt or directly undergo functional impairment, leading to death.

As the monolithic cannot undergo antifragilization, it cannot adapt to the environment. To survive, the monolithic would necessarily have to be pre-adapted to all the possible futures, or die when the unexpected comes – a point made in [Taleb, 2012].[25]

[24] "Organisms are harmed by the absence of low-level stressors (hormesis)" [Taleb, 2012]

[25] The above applies to the monolithic both on the space-dimension – no components – and on the time dimension – immortal.

Principle 15. The monolithic is fragile, and therefore comes with an expiration date.

Why is it important? No matter how strong the monolithic seems – such as a dictator or someone who fully identifies with his ego – it is bound to be broken by some stressor he cannot adapt to.

Now, let's examine the relationship between strength (the capacity to resist to stressors) and antifragility (the capacity to benefit from stressors).

Let's take two entities, A and B, with the same internal structure and the same number of homogeneous components, and let's say that A's components are stronger than B's.

The First Damage Threshold of A is higher than B's, because given two sets of components whose toughness is normally distributed with the same variance but two different averages, it is more probable that the weakest component is found in the set with the weakest average.

Similarly, A's Functional Damage Threshold is higher than B's. The number of components which need to get damaged for the entity to become functionally impaired (let's call it m) is a function of the internal structure only, which is the same for both A and B; so, they have the same m. The x delimiting the right side of the cumulative distribution of component strength underpinning a number of components equal to m is higher in A (plotting the two distributions across the same axis x indicating the stressor magnitude, A's distribution of component toughnesses is shifted to the right compared to B's). x would be the position of the First Damage Threshold.

Moreover, the distance between A's two thresholds is larger than B's, because given the same internal structure (which determines m), the value of the Functional Damage Threshold is the First Damage Threshold plus the difference between the toughnesses of the first and the mth component in ascending order of toughness. Because A's distribution is assumed to have the same shape as B's but higher average, the difference will be higher in A. Both the value of the First Damage Threshold and of the difference are higher in A, so the distance between the thresholds will be higher in A than in B.

The larger the distance between the two thresholds, the more likely an entity is to exhibit an antifragile response[26], so principle 16 can be formulated.

Principle 16. Given an internal structure, the stronger the entity, the more antifragile. The weaker, the more fragile.

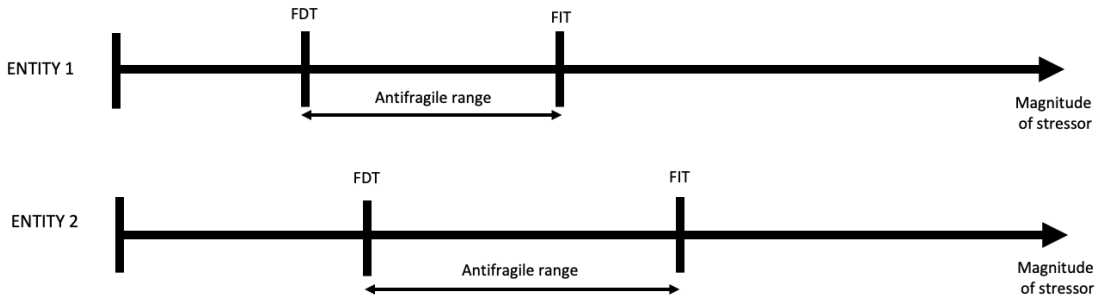

Figure 16: *Entity 2 is stronger (demonstration in the previous paragraphs).*

Why is it important? Demonstrating this principle will be necessary to justify the next one.

[26]Here, the focus is on the width of the stressor magnitude range causing an antifragile reaction, not on the width of the reaction itself. For example, a stronger person at a gym is more likely to exhibit an antifragile response because it can lift a larger range of weights benefiting from it – whereas a weaker person can only lift that much without suffering injuries. The focus is not on the fact that the weaker person can develop muscle faster, conditional on lifting the same amount.

Lack of micro-damage makes an entity fragile.[27]

[27] "Hygiene makes you fragile by denying hormesis" [Taleb, 2012]

Principle 17. Micro-damage is followed by antifragilization; lack of micro-damage is followed by fragilization.

The above follows from principles 13, 14 and 16

Antifragilization is the increase in the range of stressors which causes an antifragile response. Fragilization is its decrease.

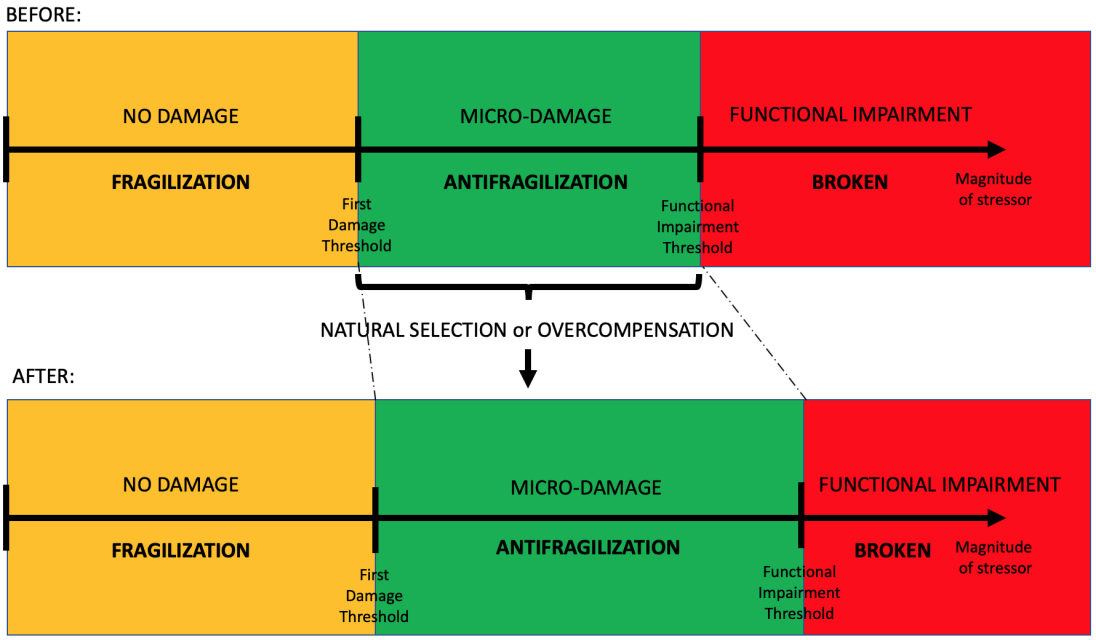

Figure 17: *Antifragilization: the distance between the First Damage Threshold and the Functional Impairment threshold increases following a stressor hitting the entity with magnitude between the two thresholds.*

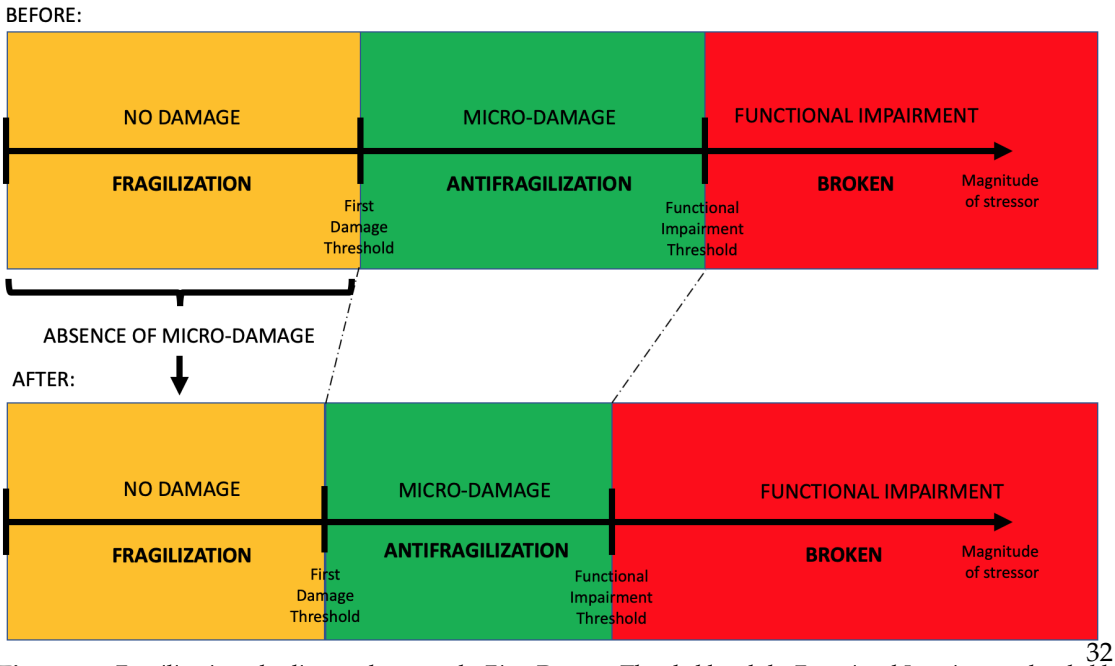

Figure 18: *Fragilization: the distance between the First Damage Threshold and the Functional Impairment threshold decreases following the lack of stressors hitting the entity with magnitude between the two thresholds.*

As already clarified, it is possible that antifragilization happens with absence of damage (for example, following placebo or nocebo effects). In other words, the cause of antifragilization is not always micro-damage. However, there are very strong associations between micro-damage and antifragilization and lack of damage and fragilization:

- Micro-damage always causes antifragilization and prolonged lack of micro-damage always causes fragilization.
- Sometimes, micro-damage and antifragilization share a common cause – for example, exposure to stressors, so that indirect pathways can arise which use exposure to stressors as the trigger for antifragilization.
- Following the previous points, it makes sense for our body to develop some system of brain-mediated inference of need for adaptation or reaction even if no physical stressor hits our body (hence the placebo and nocebo effects).

As will be clarified in section IV, damage is a strong signal for the need to adapt, hence the rationale for the two points above.

vi.1 (Anti)fragilization and scale

Due to the fact that the First Damage Threshold at a given level corresponds to the Functional Damage Threshold of the level below (principle 5), and thus the former follows the movements of the latter, the following principle can be formulated.

Principle 18. (Anti)fragilization at one level corresponds to (anti)fragilization to all levels above

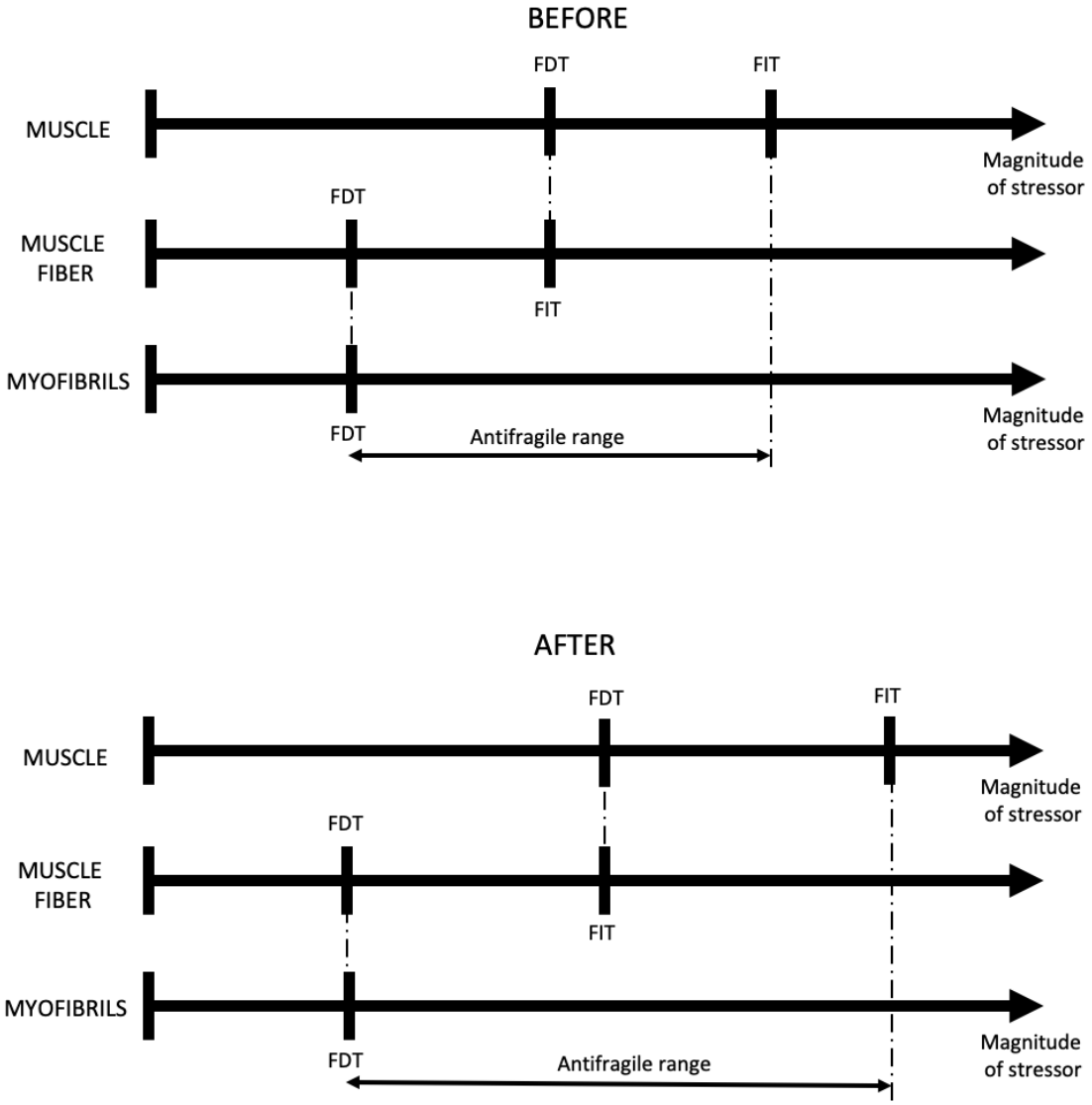

Figure 19: *Antifragilization at the muscle fibers level, causing them to become stronger, ends up affects the levels above too.*

Why is it important? In the following sections, different types of internal structures over which components organize will be discussed. It will be fundamental to have internalized how levels interact and influence each other.

vii. Equilibrium (or lack thereof)

The environment tends to generate stressors stochastically. Assuming that no stressor is strong enough to "kill" the antifragile entity, the entity adapts to them following the two rules below:

- If a stressor hits the entity above its First Damage Threshold, the entity will undergo antifragilization and the threshold itself will increase (because the entity is now stronger).
- If no stressors hits the entity above its First Damage Threshold for long enough, the entity will undergo fragilization and the threshold itself will will decrease (because the entity is now weaker).

The probability that the next stressor hits above or below the FDT depends on the position of the FDT itself, and that a stressor hitting it above pushes it higher and failure to receive hits above pushes it lower. It follows that over the long run the FDT will move around a certain position, function of the environmental distribution of stressors.

This is assuming no intelligent attempt to predict the stressor distribution by the entity.

Principle 19. An antifragile entity tends towards a stable equilibrium with the sample of stressors it recently experienced (assuming no "intelligent proactivity").

Why is it important? Adaptive entities are dynamic and the environment produces stressors stochastically; these two conditions make very difficult to predict how an adaptive entity exposed to an environment will evolve. The presence of an equilibrium allows us to estimate the direction towards which it will evolve, and will be determinant to understand the following principle.

Let's examine the expression "recently experienced" in the sentence above. The environment generates stressors stochastically. An entity only experiences a sample of the environmental distribution of the stressors, never the distribution itself. Moreover, the effect of experienced stressors adds up non-linearly and with some path-dependence because, in general, how much stronger an entity becomes after having being hit by a stressor is inversely proportional to how strong it already is (e.g., muscle weight gain as a percentage of bodyweight tends to decrease the more years one spends exercising at the gym).

As a consequence, the equilibrium of principle 19 is fragile, in the measure the distribution of stressors is fat-tailed and the "memory"[28] of the entity is short.

The more an entity adapts to a non-representative sample of the environment, the more it becomes fragile to what's outside of that sample.

Principle 20. An antifragile entity exposed to a non-representative sample of the environment achieves a partial equilibrium which is fragile and thus temporary.

Why is it important? Too often, we "protect" those we care about, limiting the environment they can experience. This makes them more fragile.

[28] What constitutes the "memory" will be described in a later section.

III. Adaptation Dynamics: Tinkering With Adaptation

This section describes how changes in internal structure, exposure and adaptation rate affect (anti)fragility and (anti)fragilization.

i. Internal structure and (anti)fragility

Imagine a small newspaper which, after a downturn in the market, lacks the revenue to keep all of his employees and has to fire one. A stressor (the market downturn) causes one of the components of the entity (an employee) to become damaged (to be let go).

One year before the market downturn, the newspaper company and its employees were already receiving some minor stressors from the market, such as indications of readers' lower engagement in some categories. Because the employees and the company viewed themselves as monolithic entities, the feedback had no chance to damage any part of them, making them adapt. The employees continued unaffected by the feedback without adapting. When the market downturn arrives and the company with his employees proves unfit for the market, something has to be damaged. In this case, it is one of the employees.

Instead, imagine that the employees viewed themselves as each comprising a multitude of behaviors. They're not monolithic anymore. Now, stressors such as market feedback can damage one of their components – one of their behaviors. For example, people not being engaged anymore by articles on theater might lead the writers to "kill" their habit to write about theater and would get them to write about, say, sports. In this case, when the market downturn arrives, because its employees already adapted, the company is fitter and is not hit anymore so strongly that it has to let go an employee.

A mismatch between an entity and the environment requires damage to adapt. The more levels of component there are in that entity, the lower will the damage materialize, possibly sparing the entity itself. In more personal words: to adapt you have to suffer, but you can choose the level at which you will suffer: your whole self, or one of your mental patterns. The lower the level, the lower the suffering.

In more technical terms, the more levels of component there are in an entity, the more it will be able to adapt, for lower-magnitude stressors will be able to cause micro-damage causing antifragilization.

(In the principle below, centralization and specialization are assumed not existing; they will be the topic of the next principle.)

Principle 21. Dividing the components of an entity to create additional levels increases antifragility and decreases brittleness.

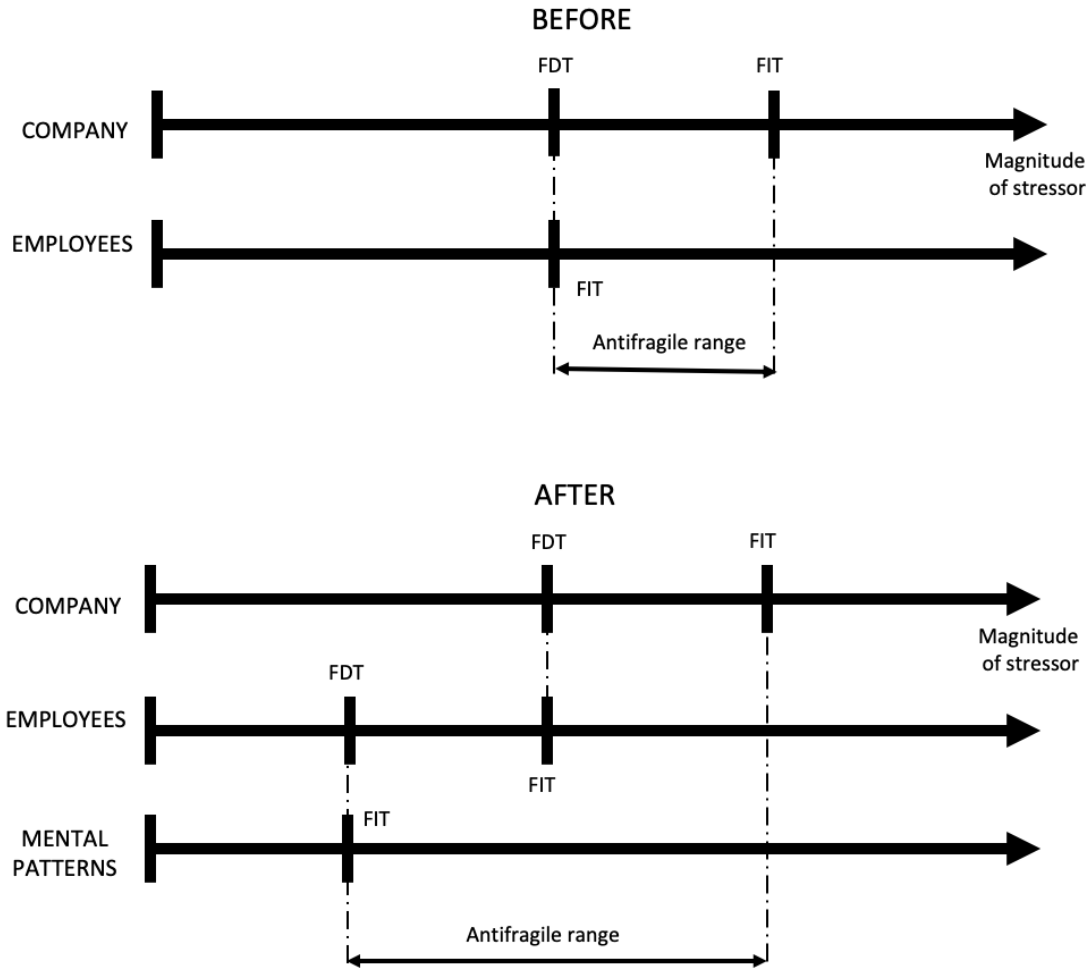

Figure 20: *Making the second level able to undergo micro-damage by creating an additional level below it, in which components can undergo independent failure causes the antifragile range to expand.*

Conversely, removing a level of components from an entity (flattening it) increases fragility.

Why is it important? In an everchanging environment, change is necessary. Adapt, or get hurt. However, we can choose the level at which damage takes place. By turning the monolithic into a container of components, we can harness lower-magnitude stressors for change.

Why is this dangerous? On the other side, if low-magnitude stressors contain information which would direct adaptation in the opposite direction compared to where high-magnitude stressors would, over-adapting to the former might mean maladapting to the latter.

The previous principle describes what happens when levels are added or removed. The following

one describes what happens when the number of components at a given level is changed, without directly adding or removing levels.

When an entity increases or decreases the number of components at one of its levels – for example, when a company grows its headcount or when a muscle grows more myofibrils – two phenomena with contrasting effects come in play:

- **If the internal structure is kept constant, increasing the number of components decreases brittleness.** This is the case of a muscle, in which growing myofibrils does not change the internal structure: muscles still contain fibers, fibers still contain myofibrils, and at any given level all components cover a similar function.
- **Increasing the number of components is often followed by a change of internal structure (centralization, specialization, or both) which increases brittleness.** This is the case of a company, in which growing headcount is usually associated with growing the number of levels in the hierarchy or in growing the number of specializations single employees have.

The relative intensity of the two effects above depend on the entity at hand. In some entities the former effect prevails, while in others, the latter does.

Principle 22. Increasing the number of components in an entity increases brittleness in the measure it leads to centralization and/or specialization, and decreases it otherwise.

Why is it important? Judging companies, countries and personal empires, size is often associated with toughness. Sometimes it is, sometimes it isn't. This principle allows to discriminate between cases.

Note: some degree of specialization and centralization can actually be beneficial, increasing overall toughness, but all specialization and centralization decrease functional or geographical redundancy adding sources of brittleness.

ii. Exposure and (anti)fragilization

The following principle completes a topic covered in principle 17, which held that micro-damage is followed by antifragilization whereas lack of micro-damage causes fragilization. Here, the focus is on clarifying to the readers the effects of of changing exposure on (anti)fragilization.

Principle 23. Increasing exposure to stressors with magnitude in the antifragile range increases antifragility.

Reducing exposure to stressors with magnitude in the antifragile range reduces antifragility.

Increasing exposure to stressors with magnitude below the First Damage Threshold has no effect (it's lack of micro-damage that causes fragilization, not stressors with magnitude lower than the First Damage Threshold). It's lack of discomfort which is problematic, not presence of comfort.

Why is it important? It explains why lifting weights and accepting constructive feedback is important.

iii. Adaptation rate and (anti)fragility

"Every plane crash makes the next one less likely, every bank crash makes the next one more likely. – [Taleb, 2012]"

The difference between the two examples in the quote above, is that in aviation, most plane crashes are the result of non-systemic errors (i.e., damage doesn't transfer quickly) whereas there are processes to disseminate the learnings of incidents investigations to the whole industry (i.e., adaptation transfers quickly). Conversely, in banks, damage tends to have systemic effects (i.e., to travel quickly) whereas learnings tend to stay inside the individual or firm producing them (i.e., adaptation is slow). The point is, the relative speed of damage respect to adaptation has a great influence on whether the reaction of an entity to damage will be fragile or antifragile.

This subsection covers the relationships between the propagation speed of damage across components and levels, and the rate at which the entity adapts to it.

Imagine an entity comprising multiple levels of components and exposed to environmental stressors. Each time a stressor hits it, if it is strong enough, it will damage components which were previously undamaged. In general, the weakest components will be damaged first, then the next-weakest ones, and so on. In brittle components, the process is better approximated by: the weakest ones first, then the ones closer to those already damaged, and so on.[29]

Multiple components can be damaged at once by a very strong stressor, or sequentially, by multiple stressors of low-to-moderate strength. For example, during a gym session, lifting a weight for the first time might break a few muscle fibers; lifting it a second time might break a few more muscle fibers, and so on.

In general, the propagation of damage will be both horizontal (the next component being damaged is at the same level of those already damaged) and vertical (a component at a higher level gets damaged if enough components at the level immediately below are damaged; for example, if enough fibers are torn, a muscle might tear).

The brittler an entity, the faster the vertical propagation speed compared to the horizontal one.

In complex entities, damage can travel vertically both bottom-up and top-down. For example, a flu might make a first employee of the post office sick, which might then make a few other employees sick (horizontal propagation), which makes the postal office unable to work properly (bottom-up propagation), which causes citizens in the district to be unable to ship and receive packages (top-down propagation).

In antifragile entities, damage propagation is offset by adaptation. In the example above, the sick employee could have been sent home to recover. The crucial factor is the relative speed of damage versus that of adaptation. If the flu propagates too fast, adaptation will be negligible until the end of the event. If damage propagates too slowly, brittleness will be negligible.

For the purposes of this subsection, the entity does not have to be antifragile to adapt. For example, if a crack is made into metal, the grains near the start of the crack will elastically bend to facilitate load-sharing.

In general, the faster the damage propagation rate (which is closely related to brittleness) compared to the adaptation rate, the more the entity is likely to exhibit a fragile response; the slower, the more the entity is likely to exhibit an antifragile response.

[29] The actual rule is: the component with the lowest ratio between load as subjectively perceived and individual toughness is the first one to break. Location, shape of the stressor impact area and other considerations can influence the numerator.

The adaptation rate is considered 0 for non-antifragile entities.

Principle 24. Entities exhibit fragility when damage propagates faster than adaptation.

Why does it matter? Policies, technologies and organizational structures that allow adaptation to propagate fast are more effective in preventing catastrophes.

Why is it dangerous? On the other hand, as explored in the next principle, an excessively fast adaptation rate can promote maladaptation (adaptation to non-representative samples of the environment). It becomes then vital to modulate adaptation speed based on the magnitude of consequences of failing to adapt.

Principle 25. The brittler an entity, the less time it has to adapt to micro-damage before suffering functional impairment.

Why is it important? Brittle entities need super-fast emergency management systems. Conversely, redundancy spent for being less brittle produces savings as slower emergency responses can be afforded and as fast-spreading harmful events will cause less damage overall.

iii.1 Damage propagation and adaptation propagation

Damage usually propagates bottom-up (myofibrils – the lowest level – break first, causing muscle fibers to break, causing muscles to tear).

Let's take an homogenous entity, made of multiple components organized in multiple levels. Let's call m_i the number of components which have to undergo functional impairment at level i for a component at level above to become functionally impaired. The lower m, the faster damages propagates vertically.

It would therefore seem that high values for m are desirable. It depends.
High m slow the vertical propagation of damage, which is good in case of high-magnitude stressors which would endanger the survival of the entity as a whole. On the other hand, shielding the upper levels of the entity from low- and medium-magnitude stressors means that those levels are more likely to undergo fragilization.

This apparent paradox – the tension between avoiding fragilization at the upper levels and avoiding extreme shocks, between avoiding fragilization at the upper levels and avoiding maladapting to non-important stressors – will be partially resolved in section VI. For the moment, the reader is invited to consider how vertical damage propagation and vertical adaptation propagation are linked and how they create opposite needs.

iii.2 Optimal adaptation rate

For the moment, let's assume a static environment: an environment whose distribution of stressor magnitudes is constant over time. The assumption will be lifted after the following principle.

Given a distribution of stressor magnitudes, the probability that the entity fragilizes or antifragilizes in response to a stressor is inversely proportional to how much it already fragilized or antifragilized,

respectively (if the entity already antifragilized a lot, the First Damage Threshold is high and stressors are more likely to hit below it; if the entity already fragilized a lot, the First Damage Threshold is low and stressors are more likely to hit above it). It follows that, given an entity with a history of stressors it has been exposed to, the most recent ones will have had a disproportionate influence on its current adapted form.

The faster the adaptation rate of the entity – the more it overcompensates stressors which caused micro-damage and the faster it reacts to a lack of micro-damage with fragilization – the more the effect described above matters.

With an approximation, it can be said that the faster the adaptation rate of an entity, the shorter its "memory" of its damage history and the more it will be adapted to recent samples of its environment only (overfitting).

Imagine a town which keeps archives of the length of the last 200 years of floods. Even if the last 30 years saw no major flood, the habitants will still be reluctant to build a warehouse too close to the river. In a town in which no such archive is kept, instead, the new generation who didn't experience any flood might be tempted to change the building policy and build the warehouse near the river, where it is vulnerable to floods.
The second town in the example above has a faster adaptation rate (it adapts based on the last 30 years of data, not on the last 200) and thus maladapts: it adapts to the recent sample of stressor, which is non-representative of the environment, and becomes more fragile to the environment as a result.

As a rule of thumb, **in a static environment, increasing adaptation rate increases fragility.**

There are many factors which contribute to a faster adaptation rate, and thus to fragility. A non-comprehensive list includes:

- **Lack of tradition.** Conservatism protects against adaptation to ephemeral trends and non-representative samples of the environment.
- **Centralization.** The more centralized is decision-making, the more the incentives to overfit to recent trends. This statement will be justified in section VI.
- **Zero-sum games.** In zero-sum games, redundancies are opportunity costs, which could be better spent elsewhere to take the upper hand in competitors. In positive-sum games, redundancies are opportunities themselves, ways to benefit from the unexpected. For example, an employee playing a zero-sum game will invest all his time in the tasks at hand, aiming for a promotion – even though promotions are a zero-sum game themselves, for usually promotions are a "only one gets it" game. Instead, an employee playing a zero-sum game will keep ample pockets of free time between meetings (say), as he is not too concerned to maximize his promotion chances. These pockets of time (redundancies) can be used to help a colleague, learn a new skill, assist a potential customer, etc. Continuing the example, the employee playing the zero-sum game will be more likely to overfit to any trend no matter how short, in his need to constantly get the upper hand on his peers. Instead, the employee playing the positive-sum game will be more thoughtful in the trends it decides to follow, preventing overoptimization and maladaptation, and providing redundancy.

Principle 26. In a static environment, increasing adaptation rate increases fragility.

Factors that increase adaptation rate (thereby increasing the risks of maladaptation) are:

- Centralization
- Lack of tradition
- Zero-sum games

Why is it important? Not all progress is good. Progress which requires assumptions, stops being progress if those assumptions can disappear.

The previous principle assumed a static environment. Now, let's dismiss this assumption, as it is not descriptive of real-world environments which are themselves made of adaptive entities and thus everchanging.

If the environment changes, having a too slow adaptation rate is just as risky as having a too high one, for both lead to being progressively more unfit to the full distribution of stressors proper of the environment. "Too high a rate of change cancels the gains from previous mutations [and other adaptations, A/N], while too slow a change leads to misfitness." [Taleb, 2019].

The charts in fig. 21 displays the fitness (i.e., adaptation to an environment) of an entity as a function of its adaptation rate.

It is not certain whether the optimal adaptation rate for an entity matches that of the environment, is slightly higher or slightly lower. Moreover, the adaptation rate of the environment is not actually a single number: the environment comprises many adaptive entities (e.g., predators, preys, etc.) and each have its own adaptation rate. However, it is certain that both a too fast and a too high adaptation rate leads to unnecessary fragility.

The concept of optimal adaptation rate is impossible to apply to real-life, for the following reasons:

- Often, there is an asymmetry between the consequences of adapting too quickly and of adapting too slowly. Therefore, it is often beneficial to overshoot or undershoot the environmental adaptation rate.
- The environment is made of multiple adaptive entities, therefore there is not a single adaptation rate for the whole environment.
- The environment is often opaque and it is impossible to accurately measure the adaptation rate.
- The environment is everchanging and so is its adaptation rate.

That said, the following principle is presented, mostly to help the reader understand how both a too-fast and a too-slow adaptation rate are problematic, and on how the environment's adaptation rate influences what is "too-fast" and "too-slow".

Principle 27. In a dynamic environment, increasing the adaptation rate of an entity up to about the adaptation rate of the environment increases antifragility, and decreases it beyond that.

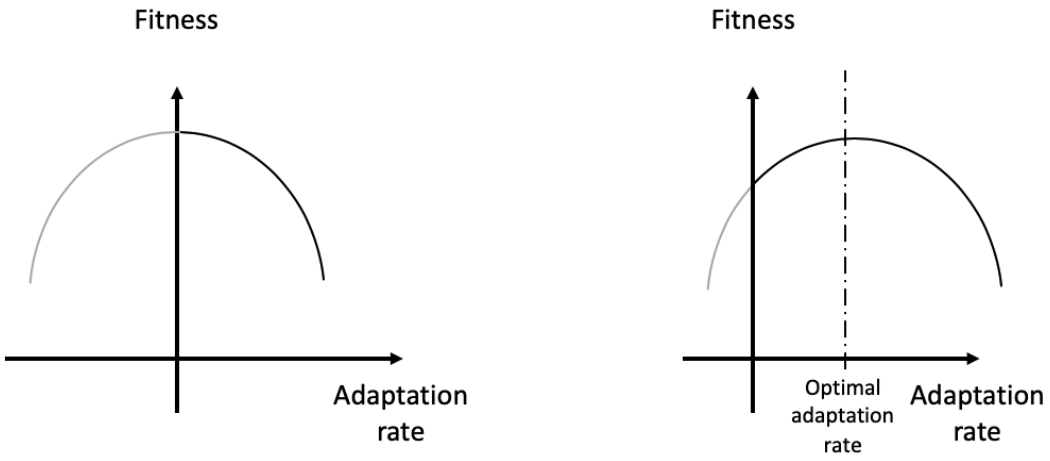

Figure 21: *Fitness as a function of the adaptation rate **of the entity**. In the first chart, the environment is static (its adaptation rate is 0). In the second chart, the environment is dynamic (positive adaptation rate). The fitness-as-a-function-of-adaptation-rate-of-the-entity curve shifts right by about the same amount of the adaptation rate of the environment.*
Note: a negative adaptation rate makes no sense; only the black part of the curve is to be considered. The grey part has been added to make easier for the reader to compare the two charts.

Why is it important? Decisions should use data sampled over the right timeframes. Too short, and maladaptation takes place. Too long, and adaptation does not keep pace with the environment.

IV. Exposure Dynamics: Why People Take Risks

i. An introduction to Risk Homeostasis

Risk homeostasis is a phenomenon in which, when given the choice between increased safety and increased performance, people choose the latter. The following case-study regarding the introduction of an automotive technology will show its relevance to everyday life.

In 1971, the Antilock Braking System (ABS) got installed for the first time as a standard on a new car model. The ABS is a device that prevents the wheels of a car from blocking when the driver brakes on a wet or slippery road. If you ever drove, suddenly braked, and felt the brake pedal pulsing under your foot, you know what the ABS is. Even though the intentions were good, the ABS proved ineffective in saving lives [of Transportation, 2009]. More precisely: it saved a lot of lives, but caused many more deaths. In fact, while in theory the ABS does increase the safety of drivers provided they drive at the same speed as they used to drive before its introduction, in practice it encourages drivers to drive faster, at the same perceived safety level. In particular, the ABS reduced the number of non-fatal crashes (those in which vehicles were going slowly) but actually *increased* the number of fatal ones (most probably, because drivers felt safe due to the ABS and now took the liberty of driving at higher speeds). By preventing the lightest incidents, it increased the perceived level of safety; the drivers responded by increasing their speed. Since high-speed incidents are much more likely to be fatal than low-speed ones, the ABS ultimately increased the overall number of deaths. In general, when people can choose between a behavioral adaptation that allows them to do something more safely or one that allows them to do something more efficiently, people tend to choose the latter. This is called risk homeostasis.

ii. The rationale for Risk Homeostasis

If the only risk incurred by people while driving were incidents and police fines, then it would be rational to drive as slow as legally possible. However, no one does that, and the reason is that incidents and police fines are not the only risks incurred by drivers.

Drivers have personal lives which extend outside their cars. A driver might have a family to come back or a job meeting to attend. A slow commute resulting in too little time spent at home or arriving late at meetings has risks (of divorce and job loss, respectively).

At any time, people are subject to hundreds of risks, and manage their resources (such as time, money, energy, reputation and focus) across all of them, taking risks in an activity when doing so would produce an increased output which could be then used to reduce other risks.

A way to visualize risk homeostasis is to imagine people as constantly trying to solve a system of equations for minimal *overall* risk, each of the equations representing an activity and linking the inputs (level of risk-taking while performing the task and the resources to spend on it) to the outputs (resources gained from the task and exposure to risks). As a non-comprehensive highly-simplified example:

- The faster I drive, the less time I spend and the more risks of incidents I have.
- The more time I have, the earlier I can arrive at work.
- The later I arrive at work, the more chances I get fired.

In the case of the example above, the driver would operate on the variable he controls (his driving behavior) to minimize overall risks (the "sum" of the risk of getting of becoming injured or jobless).

Risk homeostasis is the rational response to any situation in which an individual is exposed to multiple sources of risk and the resources to mitigate those risks are limited and shared between activities.

Principle 28. People take risks at an activity to (gain resources to) manage risks outside that activity.

These resources are only valuable for they allow us to reduce risks external to that activity.

Principle 29. People exposed to multiple risks adjust their exposure to each in order to maximize overall survival to their experienced environment.

They do so because the resources (time, money, reputation, energy, focus, etc.) they can use to mitigate those risks are limited and shared between activities.

Risk homeostasis[a] is the result.

Why is it important? This principle explains apparently irrational behaviors such as exposing one's self to survival-threatening risks when it would be possible not to.

[a] **Risk homeostasis**, the process by which a decision maker who has the option to increase the outcome of an activity while keeping its risk constant, or can decrease the risk keeping the outcome constant, will choose the former option, for activities where the risk is relatively low and all other things equal.

The placebo effect is a result of risk homeostasis. There is evidence that the animal body limits its own immune response when it is in an unsafe environment, to preserve resources and/or to maintain a more active body state which would help in case of need for a fight or flight response. Only once the animal finds itself in a favorable environment, it fully deploys its immune response [Humphrey, 2011]. The apparently irrational behavior of limiting immune response while pathogens are in the body becomes rational once other risks such as famines and predators are taken into account.

iii. How people estimate risk exposure

Mainly, people use damage to estimate risk exposure. For example, people estimate the risks of petting a dog mostly based on how often they have been bitten by a dog.

However, if people only used damage to estimate risk exposure, they would only know about risks after having been damaged by them. After all, damage only indicates "materialized risks", which is a subset of risk exposure. We want to know whether we are exposed to a risk before it gets to damage us.

Therefore, people and animals developed pain. **Pain is not – as commonly believed – a signal of damage, but a signal of vulnerability** [Dellanna, 2017]. Think about a kid who touches with his finger a hot pot in the kitchen. Milliseconds after he touches the pot, pain signals reach his nervous system and he retracts the finger. However, when he checks the finger, the finger is not damaged (provided he retracted it fast enough). Pain is not a signal of damage, but a signal of future damage if the behavior is continued – a signal of vulnerability.

Pain is, in a way, prediction of future damage; unlike many other kinds of predictions, though, it is a very reliable one, as it is based on the experience of the trillions of ancestors who lived before us, encoded in our genes. It's very Lindy[30].

Pain is a costly signal of vulnerability (it causes loss of focus, of energy and of mobility)[31]. Because it is a costly signal, it is likely to be reliable – costly and useless traits tend to disappear.

In absence of damage and pain, people use *near misses* (situations in which they could have been harmed but didn't, just because of luck; for example, a rock falling nearby without hitting anyone) *vicarious damage* (damage witnessed occurring to someone else) and *vicarious pain* (pain witnessed occurring to someone else).

However, the risk signals are not perceived as all equal. This makes sense because, as seen with risk homeostasis, people cannot decide not to take any risk. They have to take some, in order to obtain resources to mitigate other risks. In order to minimize overall risk exposure, they need to estimate exposure to individual risks as well as possible. Therefore, they need to be good at weighting risk signals appropriately.

How people do so is a very lengthy topic[32]. With a lot of simplification, it can be said that the only experiences or data we consider in our risk exposure estimations are those that cause either physical pain or strong emotions.

When people sit at a table and perform a conscious risk exposure estimation, they might also consider other data (such as statistics, etc.) but, in general, unless such data also provoked some strong emotions, their future practical behavior will disregard it. In other words, non-emotional data isn't internalized in our behavior, especially when such behavior comes with risks (skin in the game).

Principle 30. People (and adaptive entities in general) estimate risk exposure based on pain and emotional experiences.
These are lagging indicators, therefore risk exposure is mostly estimated reactively.

Why is it important? It is critical to understand that, in real-world situations in which one's experiences differ from the factual risk profile of an activity, people most often defer to the former.

Not only negative emotional experiences count; positives ones do too. This makes sense because, in general, positive emotions are associated with resources which could be used to mitigate risks (for example, receiving money which could be spent to eat) or with clues of risk mitigation (for example, the joy of getting paid, signaling that we have skills which mitigate our risks of becoming social outcasts).

Summing it up so far:

[30] Something "is Lindy" if it is antifragile and survived for long. The more something is Lindy, the less likely it is that its thriving is due to over-adaptation to a transitory trend (because the older, the more transitory trends it experienced).

[31] For this reason, pain is sometimes suppressed by other signals to allow for "performance bursts". For example, pain during a fight is often at least partially suppressed, and so is in situations where going on would bring huge rewards, such as putting our hands in that spiky bush to get those ripe berries.

[32] It will be treated in the author's upcoming book, "The Control Heuristic, Second Edition" (the topic is not "not" fully covered in the First Edition).

- As discussed when covering risk homeostasis, people engage in complex unconscious "calculations" where they weight overall risk exposure, risk exposure while performing any activity and the risk mitigation they would get by performing it, and adjust their behavior to minimize overall risk exposure.
- To estimate risk exposure and risk mitigation related to individual risks, people use pain and emotional experiences.

iii.1 The use of stressors as indicators of risk exposure

It follows from the above that in their risk exposure people mostly disregard estimations stressors that fail to cause pain or emotional experiences.

Principle 31. Only stressors that cause pain or emotional experiences are considered in risk exposure estimations.

Why is it important? Too often, we overestimate the effect that factual information might have on people's risk-taking behavior. Facts that do not cause emotional reaction are mostly disregarded in practice.

The risk model in our brain is well approximated by an antifragile entity comprising, as components, a multitude of beliefs and mental patterns regarding risks and rewards of each possible action.[33] External stressors which are strong enough to cause pain or emotions get to damage the risk model, damaging the beliefs and mental patterns which were unfit, increasing the adaptation of the risk model as a whole.

Let's take a driver who has some beliefs regarding driving, including "it is safe to drive on that road at 50mph" and "it is safe to drive on that road at 40mph". One day, while he is driving at 50mph, he almost hits a deer crossing the road. The incident was a scary experience, so the stressor (the emotion) hits the risk model of the driver and damages one of its components, the mental pattern "it is safe to drive on that road at 50mph". The next time, the driver will drive slower, as the mental pattern suggesting to drive at 50mph is not there anymore.

In the previous example, the terms "stressor", "component" and "damage" are used very broadly, but still mostly respond to the mechanics described in the previous sections of this paper.

The result is that every time we are exposed to a stressor which causes us pain or a deep emotion, the relevant mental patterns inside us which are associated with that activity get damaged (or reinforced, in the case of a positive emotion). As a consequence, our risk model adapts to the environment.

[33] The correspondence to the model presented in the first sections of this paper is not precise, as the hierarchy of beliefs and mental pattern is very loose, recursive and distributed. Moreover, the actual behaviors governing the evolution of such mental patterns depends on the mechanics of each neuron and the brain is a highly complex system. That said, the author found that the simplified model presented here has some value in helping readers understand how external stressors influence risk behavior.

Principle 32. The risk model in our mind adapts to the environment like an antifragile entity comprising a multitude of mental patterns subject to stressors whenever an external event causes pain or a deep emotion.

Why is it important? Modeling the mind as an antifragile entity made of multiple patterns allows us to understand how experiences affect future behavior.

It follows that to change one's behavior regarding an activity, one often needs to change his or her experiences with that activity (experiences being the stressors that influence behavior). For example, someone who cannot get himself to consistently go to the gym should find a way to live pleasurable experiences while going to the gym or while exercising.

iii.2 Rare stressors

Traditions, religions and genes are just some of the many systems which allow us to integrate the experience of others, to react to stressors we didn't personally experience. They are necessary to retain adaptation to high-magnitude low-frequency stressors, preventing fragilization due to the long periods in which they do not materialize.

However, some classes of stressors are so new that there has not been enough time yet for our genes and traditions to stably adapt to them. In some cases, people use the frequency of common low-magnitude stressors to estimate the frequency of the rare high-magnitude ones. For example, drivers constantly estimate how dangerous it is to drive, and in particular how dangerous it is to drive with a given style, on a given road, at a given speed, with a given weather. Because of risk homeostasis, they need granular data in order to perfectly calibrate risk taking, driving more aggressively in safer conditions and more defensively in less safe ones. Granular data means small samples, and small samples imply missing uncommon data points, such as car incidents. How can drivers, then, reactively estimate the risk of rare events, when a small sample, not large enough for the incident to manifest? Moreover, how can they estimate the risk of car incidents without having to incur them first? The solution is simple: they use the frequency of more common low-magnitude stressors such as near misses (e.g., an emergency braking or an almost hit), relying on the assumption of a quasi-constant ratio between the frequency of low-magnitude events and of high-magnitude ones.

Principle 33. People use the frequency of the more common low-magnitude stressors to estimate the frequency of the rarer high-magnitude ones.

This principle is not the expression of a voluntary behavior, but the result of fragilization in the event of a temporary lack of both low- and high-magnitude stressors hitting people's risk-model.

This is problematic when damage distributions are skewed or distorted – a point examined in the next principle.

Why is it important? Too often, policies and technologies designed with good intentions end up putting the population at danger, after taking into account how the population adapts to feeling safer, as in the ABS example which opened this section.

iv. Misdirecting risk homeostasis

The ABS example which opened this section reveals a problem with using common low-magnitude stressors to estimate the frequency of rarer high-magnitude ones: if a technology, policy, or other form of intervention protects people from just a section of a stressor distribution, it makes them underestimate the probability of stressors from the other sections of the distribution. For example, the ABS prevents drivers from experiencing some low-magnitude stressors (blocked-wheels braking events), preventing these events from damaging the mental patterns in the drivers' risk models which believe that driving at high-speed is safe. As a result, the risk models maladapt and grow the belief that driving at high-speeds is safe (no consequences and it saves time, which is itself a precious resource useful to mitigate other risks). As a consequence of the ABS' work, drivers drive faster, increasing their exposure to rare high-magnitude incidents such as collisions.

Principle 34. Protecting adaptive entities from low-magnitude stressors or shielding them from their consequences causes them to expose themselves to fat-tailed risk, decreasing their overall survival chances.

This is called the Fence Paradox.

iv.1 The Fence Paradox

The Fence Paradox [Cirillo, 2018] happens whenever a policy or technology shields people from small-magnitude stressors, causeing them to be exposed to a non-representative sample of the environment, leading them to underestimate the frequency of high-magnitude stressors and thus to underestimate their exposure to a given risk. Because people seek to minimize their overall risk exposure, the less risky an activity appears more risks they take (risk homeostasis), if that would allow them to grab increased rewards.
The increased risk taking in an activity whose risk exposure has been underestimated causes an increase to the actor's overall risk exposure.

In other words, policies and technologies which prevents people from being affected by small-magnitude stressors decreases their overall survival.

iv.2 The pain threshold

The previous principle explained why technologies and policies that protect people from low-magnitude stressors without equally protecting them from high-magnitude ones endangers them. In some cases, however, instead of protecting people from stressors or from their consequences, they limit the perception of consequences, by reducing the pain (or negative emotion) they cause. Painkillers are one such example.

Increasing the pain threshold is equivalent to increasing the First Damage Threshold: the next stressor from a given distribution will be more likely to have fragilizing effects.

Consequently, increasing the pain threshold makes an entity harder but brittler (as the First Damage Threshold increases, the Functional Impairment Threshold does not move, and thus the distance between the two decreases). Moreover, because it makes people take more risks (because less low-magnitude stressors are perceived as damaging), once the thresholds are normalized for

the increased risk exposition, the final result is that the entity is just as strong as before but way brittler and thus more fragile.

Principle 35. Suppressing pain from stressors causes an entity to take unnecessary risks and makes it more fragile, thereby decreasing its overall survival rate.

Suppressing pain is only acceptable when the distribution of stressor magnitudes is bounded (so that there is no fat-tail-stressor to underestimate) or when a more imminent risk is present and pain would impair resistance to such risk (for example, it is good not to feel pain at the legs while running away from a predator).

Why is it important? Too often, people take painkillers – medicinal or metaphorical ones – thinking it would mitigate their problems. Often, it deepens them instead, by hiding the need to change the root cause for the pain.

Pain is a signal of vulnerability: as such, in most cases, it should be acknowledged, not ignored. (Of course, there are cases in which pain is not linked to a maladaptive behavior, and suppressing it might be positive and compassionate.)

iv.3 Addiction

In addition to the points mentioned above, there is a major source of risk exposure miscalculation: "addiction to rewards", a broad umbrella of phenomena going on in our brain and hacked by a lot of modern creations. Their treatise would require doubling the size of this manuscript, so they've been taken out of scope. [34]

v. Redundancy and risk exposure

An entity which is conscious of its brittleness will tend to limit its exposure to risks, decreasing the average intensity of the distribution of stressor magnitudes it willfully exposes to, leading to longer periods with no stressor hitting above the First Damage Threshold, and thus fragilization.

As the entity gets weaker, it will find its environment more dangerous (as it is weaker) every time it fails to limit its exposure to it. For example, a weak person might limit its needs for physical strength but it cannot avoid exposure to all stressors (for example, it might get hit by a bicycle or get into a fight). The weaker the entity, therefore, the more energy it will have to spend to try to control its environment, entering a vicious circle where it becomes even weaker (as any available resource is used to control its environment rather than growing) and fragile (as the parts of its environment it can control contain lower-magnitude stressors, but the part of its environment it cannot control are just as absolutely as strong as before and relatively even stronger).

Conversely, an entity which has and can afford redundancy is less brittle and behaves the opposite way. As it can afford exposure to stressors, it exposes itself to them, undergoing antifragilization and becoming stronger, and thus more redundant, entering a virtuous circle of strengthening.

[34] They will be object of the author's upcoming book, "The Control Heuristic, Second Edition" (they're not fully covered in the First Edition).

Principle 36. Redundancy is a prerequisite to willfully and knowingly increase risk-taking in unbounded environments.

Sentient brittle entities decrease exposure to stressors in the measure they are conscious of their brittleness and decrease it in the measure they are not.

Why is it important? Redundancy is important not only for its first-order effects (damage absorption) but for its second-order ones too (propensity to expose one's self to antifragilizing stressors).

V. Ergodicity: How to Reap The Gains of Antifragility

Ergodicity is one of the most talked topics in Behavioral Economics of late, after it has been shown to explain apparently irrational decision making in gambles [Peters, 2009]. It might seem unnecessary to bring it up in this paper on damage and adaptation. However, adaptation requires being able to play repeated games, a property linked to ergodicity. The link between antifragility and ergodicity will be further discussed in section ii. Before that, an introduction to the concept.

i. An introduction to ergodicity

An activity is ergodic if the outcome of it being performed by a single actor n times coincides with the outcome of n actors performing it once. For example, in Russian Roulette, the expected outcome of $n \to \infty$ players playing it once is $\frac{5}{6}$ (i.e., the survival rate) times the reward for playing it once, whereas the likely outcome for a single player playing it $n \to \infty$ times is death. Because the two outcomes differ, Russian Roulette is a non-ergodic activity for a Russian Roulette player.

In excessively simplified terms, it could be said that activities are non-ergodic when they present a risk of "game-over."

Non-ergodicity is caused by the presence of absorbing barriers – conditions which, if crossed by a player's path, prevent him from moving away. Examples of absorbing barriers are death for people, bankruptcies for companies and loss of licenses for professionals.

Ergodicity is the property of an actor, an activity, and the exposure of the former to the latter. Continuing the example above, Russian Roulette would almost[35] be an ergodic activity for a player who points the gun to a doll's head rather than to his own, or for a company which hires Russian Roulette players . Who is performing an activity and how he is exposed to it matter.

Principle 37. Ergodicity is the property of an actor, an environment (comprising one or more absorbing states), and the exposure of the former to the latter.

Why is it important? Considering ergodicity only as a function of an activity prevents us from thinking about how actors can change their exposure and how they can change themselves to increase their chances of survival.

[35] "Almost" because a player could run out of dolls, or a company could run out of players to hire.

i.1 Morality and ergodicity

Morality largely maps over ergodicity[36], in the measure in which most civilization consider immoral actions which tend to bring non- ergodicity to individuals or to populations. For example, the following actions are largely considered immoral: killing, stealing (lowering one's resources, so that it is more difficult for him to survive), and selfishness (avoiding to contribute resources to a shared pool, which has been shown to increase ergodicity [Peters and Adamou, 2015]).

Conversely, virtues and heroism are largely associated with acts which decrease the possibility that the community hits an absorbing barrier.

Moreover, when ergodicity is present and thus repeated games are not only possible but also expected, malicious behavior is discouraged (because the selfish actor will have to face retaliation from the people it wronged during the next iteration). Therefore, ergodicity encourages pro-social behaviors showed to also have economic benefit over the long run, such as sharing resources [Peters and Adamou, 2015].

Principle 38. Most Lindy moral rules map over avoiding to expose members of the community to absorbing states.

Why is it important? It shows how important is ergodicity to our survival and how deeply rooted is the concept inside us, even if we do not know its technical formulation.

ii. Ergodicity and antifragility

ii.1 Non-binary ergodicity

The most-accepted definition of ergodicity is binary: either an activity is ergodic (time-average and expected value coincide) or it's not. However, few activities in the real world outside of games are ergodic. If people only limited themselves to ergodic activities, they wouldn't get much done.

Instead, people are more interested in the question "how much ergodic is this activity?", so that they can compare activities, take into account costs, risks, rewards and probabilities of performing the activity again and adjust their behavior accordingly, as described in the section on risk homeostasis.

If actors had infinite life expectancy, then only a binary definition of ergodicity makes sense. However, the life expectancy of actors is limited, so it is possible that for them, during their life-time, time-average is likely to coincide with expected value for the activities they participated to which with little chances of "game over".

The author knows that the previous sentence is not technically correct, and yet, it makes perfect sense from an individual actor's point of view. Driving a car is a non-ergodic activity, but sometimes we drive as if it were, when facing other activities which are way less ergodic. Risk homeostasis dictates that risks and ergodicity are evaluated subjectively and relatively to one's life expectancy, not absolutely.

[36] At least the most Lindy and antique forms of morality – not thought experiments from modern academics. (Of course Lindy moral rules are about avoiding absorbing barriers; otherwise, they would not have survived.)

Principle 39. Ergodicity can be considered as non-binary when people live in an environment with multiple risks and limited resources to mitigate them which are shared between risks, so that mitigating one risk implies increasing exposure to another.

Why is it important? In the real world, people constantly unconsciously evaluate "how much an activity is ergodic", even though technically the answer would be "few activities outside of games are ergodic."

ii.2 Efficiency and survival

Efficiency is usually expressed as the ratio between output and input. It is often used to answer questions such as: how much of the input contributed to the output?

Within a short observation period, the difference between output and input is called redundancy or waste.

However, as the observation period gets longer, more and more of what was previously considered a redundancy is used to mitigate some extreme unexpected environmental condition.

The longer the observation period, the more likely it is that something which was previously considered efficient breaks, whereas something which was considered inefficient survives.

As a rule of thumb, inefficiency coming from redundancy translates to survival over a long enough timeframe, whereas inefficiency coming from "burning" redundancy goes to waste.

Principle 40. The shorter the expected lifetime and the more likely that absorption barriers come from peers (competition), the more efficiency is important to survival. The longer the expected lifetime and the more likely that absorption barriers come from the environment, the more redundancies translate to survival.

Efficiency is performance under the conditions that materialized.

Efficiency is always related to a result, and restricting the definition of the result or the dimensions over which it is measured (including time) is the fastest way to increase efficiency and to decrease survival.

Efficiency is negatively correlated with redundancy, which is itself negatively correlated with brittleness. So, the more efficient, the brittler.

Why is it important? People and companies expecting long lifetimes should focus less on efficiency and narrow definitions of "results" and more on inefficiency (redundancies) and broad definitions of "results".

ii.3 The level-dependency of ergodicity

For entities composed of multiple levels of components, ergodicity is evaluated differently at different levels.

xiii. Other principles

The principles listed in this paper are non-exhaustive. There are many other considerations to make; here, the author presented the ones the most closely related to damage, the topic of this manuscript.

Other important principles can be found in [Taleb, 2019].

Conclusions

Policies, technologies and organizational systems should be judged not (only) for what they do, but for how they cause populations to adapt to their presence.

To understand adaptation, one must understand how damage propagates and how vulnerability to damage is perceived. Hopefully, this paper will have provided the readers with some tools to do that.

The topic of adaptation is extremely complex, as the interactions within and between adaptive entities are themselves complex. Therefore, this paper cannot be considered exhaustive; it is only designed as an introduction to a lengthy but passionating topic.

Acknowledgements

The contents of this paper would never have occurred to the author without him having read the works of Nassim Nicholas Taleb, Joe Norman, Ole Peters, Pasquale Cirillo, Rory Sutherland and many others including the ones listed below.

Glossary

component components are the constituents of entities which can be damaged independently. For example, rocks are made of crystals, humans are made of bones and muscles, muscles are made of fibers, and companies are made of employees. In case of entities made of multiple levels of components (e.g. man → muscles → muscle fibers), "the components of an entity" refer to the components at the level immediately below the entity, unless otherwise specified. For example, "the components of muscles" refers to its muscle fibers, not to its cells. 1

damage any permanent or semi-permanent modification to an entity or component causing it or any of its components to undergo functional failure. For example, both single muscle fibers torn and full muscle tears are forms of damage from the point of view of the person. Glass cracks and plastic deformations are also forms of damage, whereas elastic deformations aren't – they can immediately be reversed). 1, 13

entity any being which can "break" (for living beings, this means dying). For example, humans, bacteria, teacups and civilizations are all *entities*. 1

Fatigue The process of accumulating damage at the component level because of an insufficient healing rate (if present at all). 1

First Damage Threshold the magnitude of the weakest stressor able to cause one of the components of the entity being hit to become functionally impaired. 1, 26

functional failure a component is said to undergo "Functional Failure" if it cannot contribute anymore to the ability of the entity it belongs to to resist stressors. For example, a tore muscle is under functional failure until healed (it cannot contribute anymore to the survival of the person it belongs to). Functional failure can be permanent (a broken wall) or temporary (a torn muscle). 1

Functional Impairment Threshold the magnitude of the weakest stressor able to cause functional impairment to the entity being hit. 1, 26

micro-damage damage causing functional failure to a component but not to an entity. Functional failure of a component always corresponds to (at least) micro-damage from the point of view of the level the component belongs to. For example, a single torn muscle fiber represents functional failure for the muscle fiber but is micro-damage from the point of view of the muscle. 1

risk homeostasis the process by which a decision maker who has the option to increase the outcome of an activity while keeping its risk constant, or can decrease the risk keeping the outcome constant, will choose the former option, for activities where the risk is relatively low and all other things equal. 1, 46

stressor any event able to cause damage to an entity, if strong enough. Punches, famines, wars and financial crises are all examples of stressors (the first three can cause damage to humans and the last two can cause harm to companies). 1, 9

References

[Cirillo, 2018] Cirillo, P. U. o. D. (2018). Of risk, fences and unavoidable falls. .

[Dellanna, 2017] Dellanna, L. (2017). Pain is a Signal of Vulnerability. .

[Humphrey, 2011] Humphrey, N. (2011). The evolved self-management system. Edge.com .

[of Transportation, 2009] of Transportation, U. D. (2009). The Long-Term Effect of ABS in Passenger Cars and LTVs.

[Peters, 2009] Peters, O. (2009). Optimal leverage from non-ergodicity. Papers 0902.2965 arXiv.org.

[Peters and Adamou, 2015] Peters, O. and Adamou, A. (2015). An evolutionary advantage of cooperation. Papers 1506.03414 arXiv.org.

[Sutherland, 2019] Sutherland, R. (2019). Plumbers always have the best restaurant recommendations. The Spectator .

[Taleb, 2012] Taleb, N. N. (2012). Antifragile: Things That Gain From Disorder.

[Taleb, 2019] Taleb, N. N. (2019). Principia Politica. .

[Taleb et al., 2014] Taleb, N. N., Read, R., Douady, R., Norman, J. N. and Bar-Yam, Y. (2014). The Precautionary Principle (with Application to the Genetic Modification of Organisms). .

[Thorp, 1975] Thorp, E. O. (1975). Portfolio Choice and the Kelly Criterion.