Visualizing Biological Evolution

Guided by Prof. Venkatesh Rajamanickam

Nishita Nirmal

196330007

M.Des Interaction Design (2019-2021)

IDC School of Design, IIT Bombay

Contents

Introduction

The Theory of evolution is widely known to be a means of explaining our origins and how we came to be. It is also often viewed as a hypothetical mechanism that becomes visible in retrospect over millions of years. While this is partly true, in recent years, it has become possible for scientists to observe and record evolution live, in the wild, as demonstrated by Peter and Rosemary Grant in their 40 year Finch Study (Grant & Grant, 2014). The field of Evolutionary Biology is continuously updating itself, and with it, our collective understanding of evolution.

In his book, 'The Beak of the Finch, ' Jonathan Weiner writes, "And our power to drive Darwin's process, like the power of the process itself, is not hypothetical. The industrial revolution was changing environments and, with them, the course of evolution even before Darwin published the Origin." (Weiner, 1994)
He describes the human species as an ecological dominant, driving and accelerating evolution, both in the wild and within our backyards.

While we are influencing the evolution of species on this planet in multiple ways, i.e., through domestication, which leads to the hybridization of species and through the effect of human-made activities on the climate, evolution also has direct and immediate implications for our Agricultural and Medical sectors. Both the fields of Agriculture and Medicine battle organisms that are continually evolving and resisting us, i.e., Insects and viruses, respectively. As pesticide companies build more robust and stronger insecticides, they fail to understand the evolutionary mechanisms that allow insects to adapt and evolve against the selection pressure applied by these pesticides. The same cycle can be identified in the production of antibiotics at big Pharma companies. Ecologists Robert May and Andrew Dobson say that this "may help to show that evolution is not some scholarly abstraction, but rather is a reality that has undermined, and will continue to undermine, any control program that fails to take account of evolutionary processes." (May & Dobson, 1986)

In his article 'Public Understanding of Science' published on the APA website, Dr. Steven J. Breckler writes, "With scientific literacy comes better decision-making capabilities and easier navigation through life's everyday challenges...Public understanding of science is also essential for the health of the scientific enterprise itself."

In this project, I aim to use Information Visualization and Data Visualization to communicate ideas from Evolutionary biology, as expressed in two well-known books about the 40-year-long Finch Study. The following sections explain the Study that this project is based on and then talk about the reference books, audience, intent of the visualization, key-questions it should answer, and the datasets used. The next part will explain the design methods and devices used, the medium, narrative structure, and visualization design I have explored. The report will end with my reflections on the process and plan for the next stage.

About the Project

Primary Goal
The primary objective of this project it to communicate the concepts of the 40 year Finch study to a general audience in an effective and engaging manner. For this purpose, I have used data and concepts from the Finch study and visualized them using design principles as guidelines.

Non-goal
Though it is an interesting direction to take, I have decided not to focus on the ‘novelty’ of the visualizations, but instead keep it as second priority.

Positioning
This project is positioned as a gateway for general, interested audiences to know more about evolution and perhaps go on to reading more about it in the future. It is not in a position to augment or redesign the original scientific visualizations of the study or the books.

Evaluation
Since the project focus is on translating and communicating the information well, I went forward keeping this evaluation criteria in mind:

1. Clarity of the final output - concepts and visualizations
2. Exploration and understanding of the Subject (Data Visualization) and Domain (Evolutionary study)
3. Design processes followed to come up with the design
4. Aesthetic usability of the output

PHASE 1 - Exploration and Planning

About the Grants' Study

This project is based on the Finch study by the evolutionary biologist couple, Peter and Rosemary Grant. They conducted the landmark study on the Galapagos islands, the same islands where Darwin first came across the finch specimens that would lead him to evolution theory. The Grants and their team conducted the study on the Island 'Daphne Major' for around 40 years starting in 1973 and collected data on the four major Finch populations. As they analyzed the data on beak and body size, mating behavior, and feeding habits of the birds, they found insights that would significantly change our understanding of evolution.

This study is quite notable because it is the first extensively recorded study of evolution in the wild. Before, most studies on evolution focused on retrospective analysis, using information on contemporary populations to test the assumptions of historical hypotheses. This study focuses on prospective analysis, studying a population through time, and recording insights about the process of speciation. It follows up on Darwin's theory of evolution, provides supporting evidence, and contributes new findings. The goal was to find out how competition for food and space, triggered by environmental changes, might affect morphology, behavior, and, ultimately, the evolution of species. If Darwin's finches were the macrocosm, they studied the microcosm on Daphne Major.

Some of the high-level findings of this study are:

1. *Evolution is more visible as the granularity of time increases. It is a rapid process.* This finding opposes the general belief that evolution can only be observed at a high-level over a long period.

2. *Selection is constantly changing.* Nature will select for opposing traits at different times. There is no ideal beak or body size for the Finch population.

3. *Selection oscillates in direction.* It can be compared to a pendulum swinging from one extreme to another, always in the act of correcting and recorrecting populations.

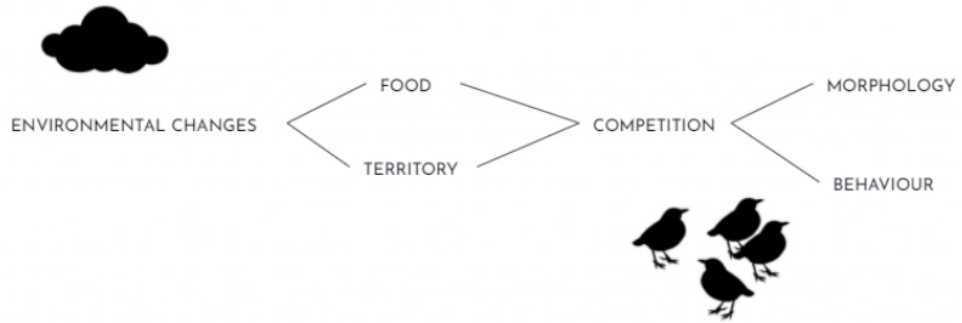


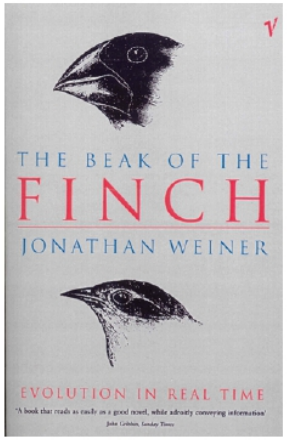
Figure 1: My illustration of the goal of the study, i.e.,to see how competition for food and space, triggered by environmental changes, might affect morphology, behavior, and evolution of species.

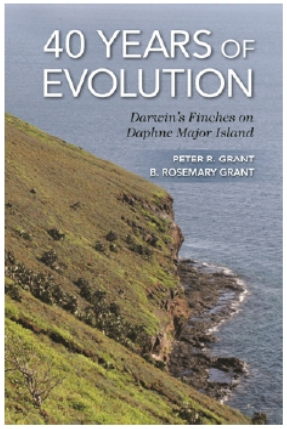
Primary Reference books

For visualization content, I am referring to two books:

1. ‘The Beak of the Finch’ (1994, 2014) by Jonathan Weiner
The first publication of this book won the 1995 Pulitzer for ‘General nonfiction.’ It outlines the Grants’ study’s key findings and weaves a narrative through time, starting with Darwin’s Theory and juxtaposing it with other studies in the field. It starts with the Finch study, then diverges to various other case studies, providing a holistic summary of a lot of the work in the domain of evolutionary biology until 1994. The book uses simple language and is meant for a general audience - casual readers and non-specialists interested in the domain.

2. ‘40 years of Evolution’ (2014) by Peter and Rosemary Grant
This book, written by the original authors of the study, is meant for students, specialists, and enthusiasts interested in evolutionary biology. It comes with the complete dataset of the study, and this book will be the primary source of data for the project. The book begins with an overview of the state of evolutionary theory before the study. It then converges into the experiment’s details, all the data collected, its statistical analysis, and its scientific implications. It does a good job of introducing scientific terminology according to the knowledge level of the reader.





(above) The Beak of the Finch (1994) by Jonathan Weiner
(below) 40 years of Evolution (2014) by Peter and Rosemary Grant

Figure 2: Overview and contents of 'The Beak of the Finch' by Jonathan Weiner

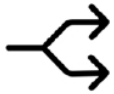
1

The Beak of the Finch (1994, 2014)
Jonathan Weiner

Audience : General Public

Provides Larger context
- Comparison with other studies in the field
- Comparison to Darwin's ideas
- Outlines the Importance of studying Evolution

Structured as a narrative
- High Literary quality
- Simple language, no stats

Diverge 

CONTENTS

Figure 3: Overview and contents of '40 years of Evolution' by Peter and Rosemary Grant

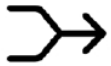
2

40 years of Evolution (2014)
Peter and Rosemary Grant

Audience : Students, specialists in Evolutionary Biology

Details the Study
- Describes the experiment setup
- Provides the Dataset and analysis
- Contains a variety of Visualisations
- Outlines Scientific implications

Concept-driven structure
- Outlines concepts systematically
- Uses scientific terminology

Converge 

CONTENTS

Audience

The purpose of defining a target audience would be to enable design for a particular knowledge-level, to fulfill specific needs, find the right platform for consumption of the visualization, and to be able to validate the effectiveness of the design. To this end, while talking about 'Domain situation' Tamara Munzner explains, "The group of target users might be as narrowly defined as a handful of people working at a specific company, or as broadly defined as anybody who does scientific research." (Munzner, 2015, p. 94)

The target audience that I have in mind is the General Public interested in learning about/exploring the concept of evolution.

Knowledge-level: Non-specialised, basic school-level exposure to the theory of evolution.

Language: Can read and understand English

Platforms: Know how to use desktop computers or laptops.

Habits:
1. Used to reading
2. Like to browse the internet

Assumed needs :
1. To discover high-level, non-specialized insights about the process of evolution
2. To develop a 'sense' of the subject matter
3. To find the visualization engaging

Intent of the Visualization

In 'Interactive visual explainers: a simple classification' Maish Nichani and Venkatesh Rajamanickam explain how the 'intent' of visualizations is presented in 4 different ways:

Instructives - explain step by step how things work. Here, the objective is to explain by enabling the reader to step through the intent sequentially.

Exploratives - offer the reader an opportunity to explore and discover the intent.

Simulatives - allow the reader to experience the intent (usually a real-world phenomenon).

Narratives - The objective is to explain by giving the reader a vicarious experience of the intent through a story. Narrative information graphics are characterized by a story told with a distinct point of view.

(Nichani and Rajamanickam, 2003)

Based on the content I am presenting and the chosen target audience, the best presentation of data for this project would be the 'Narrative' type. While there will be elements of the other types within the narrative, here is why they are not the primary choice:

Why not instructive: While this is a good option for the chosen target audience, the needs I have assumed are 'to get a sense of the subject matter' and 'to be engaged in the visualization.' To explain the 'how' of technically challenging matters like 'Heritable Variation' or 'Introgressive Hybridization' in the study in a step-by-step manner would require the audience to go through a bit of a learning curve. It might also require expertise in handling/preserving the subject's accuracy, that I do not have.

Why not Explorative: To be able to explore the data and come up with insights on their own, users would require some amount of familiarity with the domain. The audience chosen is not expected to have any specialized knowledge. However, there is some light exploration through interactive charts.

Why not simulative: The content, i.e., the data of the study, recall specific past events, and so to be able to simulate it in a way that the user can make 'choices' would require good predictive algorithms. It also runs the risk of being scientifically inaccurate.

Why Narrative: Narratives have the advantage of being inherently driven by engagement. They are also an excellent way to communicate high-level insights and chronology of events using simple language. In this case, a narrative would act as a framework for viewers to understand the chronology of the data collection and analysis, and come up with questions about how the data might have changed, had the narrative changed.

Key Questions

To develop a narrative structure, I outlined a few key domain questions that the visualization narrative should answer. They are based on the chapters of the book ‘40 Years of Evolution’.

What is a species?
C1) Speciation, Adaptive Radiation

What is variation?
C2) Finch a question of size

What is genetically heritable, and what is not? Genetic variation vs. behavioral variation
C3) Heritable Variation

What does nature select for? What causes change? How often does selection occur, and how? How is Natural selection different from evolution? What are some examples of selection events?
C4) Natural Selection and evolution. Sec. ‘ selection occurs repeatedly’, sec. ‘Selection oscillates,’ box 4.2, La Nina. El Nino.

What is fitness? What are the components of biological success? How does ecology affect longevity of life? How is mating behavior affected by ecology?
C5) Breeding Ecology and fitness

How does introducing a new species affect the ecology of the island? What role does chance play?
C6) A potential competitor arrives on Daphne Major

What is competition? What is character displacement? How does competition influence evolution?
c7) Competition and character displacement

Natural selection vs sexual selection
Misc) The beak of the Finch

I found the last few chapters of the book ‘40 years of evolution’ to be a little challenging to understand as they deal with Genetics. For those reasons, I haven’t added the Genetic aspects of Evolution in this Data story.

Data set and Existing Visualizations

The complete Dataset from the Finch study can be found at this link:
https://datadryad.org/stash/dataset/doi:10.5061/dryad.g6g3h

The details are:
Grant, Peter R.
Grant, B. Rosemary
Publication date: August 28, 2013
Publisher: Dryad
https://doi.org/10.5061/dryad.g6g3h

Dataset contains **82 files**
Dataset type: **Tables**
Attribute types: **Categorical, Quantitative**
Dataset Availability: **Static files**

This dataset is mapped to the **16 chapters** of the book '40 years of Evolution'. Out of the 82 files of the dataset, my task is to choose which files will be useful in the narrative.

I will present and explain a few of the data files and accompanying visualizations from the book in the following pages. **All the Data and Illustrations shown below are from the book '40 years of Evolution'.**

Chp.1 Speciation, Adaptive Radiation & Evolution

The study deals mainly with 4 species of ground finches -
Fuliginosa, Fortis, Magnirostris - Granivores - they mainly differ in body and beak size and are a gradation as can be seen in the density ellipses of the accompanying visualization.
Scandens - cactus specialist - has a longer and narrower beak (shown in red)

Name: **Fig. B1-02**
Dataset: **4 attributes, 54 items**
Attribute types: **Categorical, Quantitative**

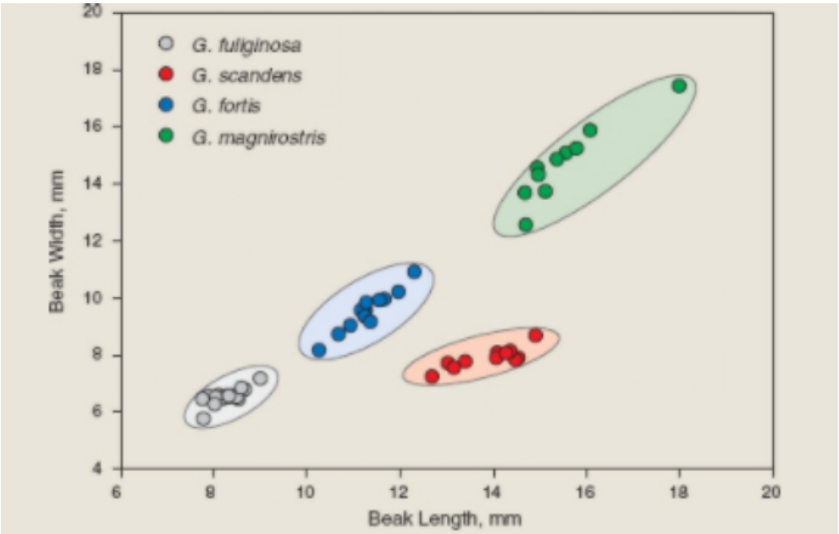


Fig. B.1.2 Morphological variation among four species of Darwin's ground finches (males) on several islands. Data are taken from Grant et al. 1985.

Island	Species	Beak length	Beak width
Espa–ola	fuliginosa	8.39	6.58
Gardner E	fuliginosa	8.3	6.58
Floreana	fuliginosa	8.33	6.54
Gardner F	fuliginosa	8.66	6.8
Enderby	fuliginosa	7.89	6.57
S. Crist—bal	fuliginosa	8.53	6.46
Santa Fe	fuliginosa	8.49	6.53
Santa Cruz	fuliginosa	8.18	6.48
Baltra	fuliginosa	8.29	6.57

Chp.1 Speciation, Adaptive Radiation & Evolution

Fig.1.1 is a Concept Diagram explaining the process of speciation. There is no quantitative data that has been used and the image purely represents a concept.

According to David Lack, three factors that affect speciation are:
1. Natural Selection
2. Diversification on separate islands
3. Competition for food

The diagram explains how a population of birds migrates to a new island (in this case San Cristobal) and colonizes it. A few birds leave the original population and disperse to another island, where, due to a new environment, the population goes through some evolutionary changes. Parts of this new population move from island to island, diverging in different ways till some of them fly back to the original island. This diverged population then competes with the original to further drive them apart.

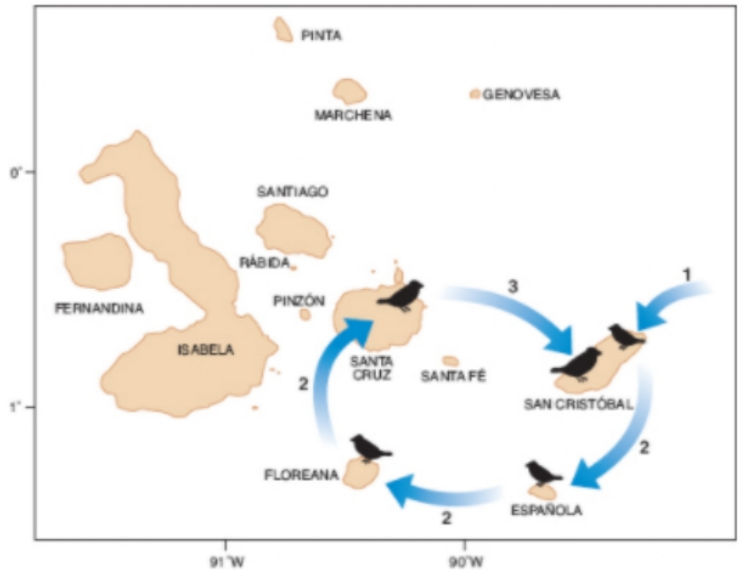


Fig. 1.1 Allopatric speciation in three stages: initial colonization (1), establishment of a second and additional populations (2), and secondary contact between two divergent populations (3). Choice of islands is arbitrary. Repetition of stages 2 and 3 in other parts of the archipelago gives rise to more species. From Grant 1981a, Grant and Grant 2008a.

Chp.1 Speciation, Adaptive Radiation & Evolution

Fig.1.2 is a **Concept Diagram** explaining the process of speciation. There is no quantitative data that has been used and the image purely represents a concept.

Species and Speciation:
What prevents species from interbreeding and fusing into a single population? What makes Darwin's finch species?
Size, proportions, plumage, morphology, song, mating behaviour

Mating:
Two populations can be called different species when they are Reproductively isolated. Between species interbreeding is rare and inhibited - they either don't interbreed (behavioural) or if they do their offspring will be infertile (genetic)

The Diagram shows the stages in the process of divergence.

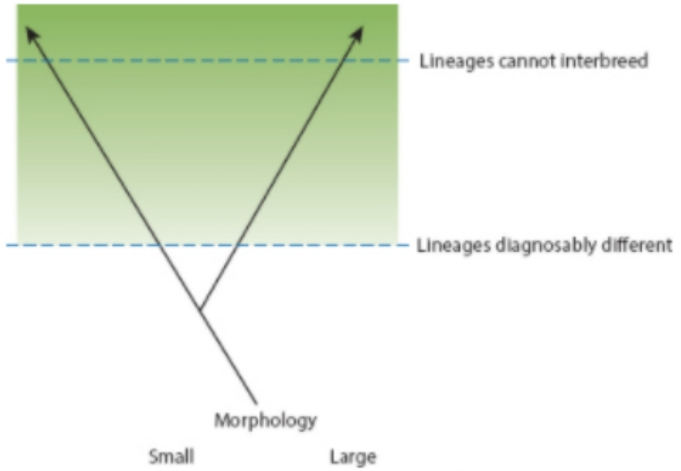


Fig. 1.2 Y diagram of speciation. The process is represented as a splitting and divergence of lineages. Opinions vary on when the lineages merit recognition as two species because divergence is gradual (discussed in Grant and Grant 2008a).

Chp.1 Speciation, Adaptive Radiation & Evolution

Name: **Fig. 01-06, 01-07**
Dataset: **8 attributes, 41 items**
Attribute types: **Categorical, Quantitative**

Fig. 01.06 and 01.07 show the **Evolutionary Trajectory of Fortis and Scandens beak sizes over 40 years**
The grey lines mark the projected 95% confidence intervals of the predicted evolutionary trajectory of beak sizes. In red, the 'Scandens' beak sizes can be seen following more or less the same path with a few minor excursions.
Below in blue is the 'Fortis' trajectory and it is visible that there was a sudden rise in mean beak size of the population around 1977 and a steady drop starting in 1984.

What this chart doesn't show is that two major events could have affected the Fortis population during this time:
1. The arrival of a hybrid on Daphne Major
2. The expansion of the Magnirostris population

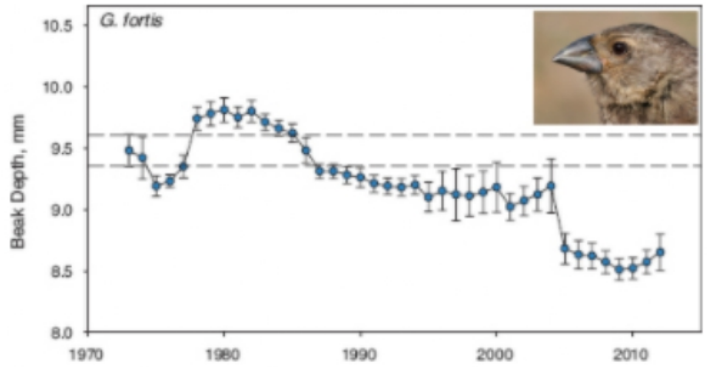


Fig. 1.6 Evolutionary trajectory of *fortis* beak size over 40 years. Means and 95% confidence limits are shown for all birds alive in each year. Parallel horizontal lines mark the upper and lower 95% confidence limits on the first estimate of a mean based on a large sample size ($n = 221$) in 1973.

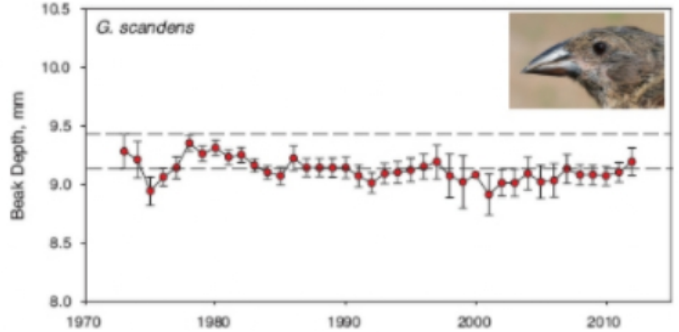


Fig. 1.7 Evolutionary trajectory of *scandens* beak size over 40 years. Means and 95% confidence limits are shown for all birds alive in each year. Parallel horizontal lines mark the upper and lower 95% confidence limits on the estimate of the mean in 1973 ($n = 71$).

Chp.2. Daphne Finches

Island and experiment setup

Fig. 2.4figure shows 4 **schematic maps** of the island. It shows how the Island 'Daphne Major' was gradually isolated from the mainland over 22,000 years.

Fig. 2.5. Is also a schematic map explaining how the Grants divided the island into parts for sampling the seeds and vegetation all over the island.

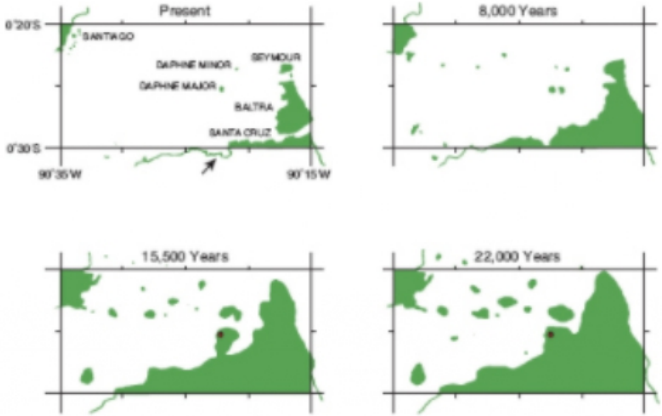


Fig. 2.4 Isolation of Daphne as a result of a rise in sea level associated with melting of glacial ice. From Grant and Grant 1996a, 2008a. The position of Borrero Bay is indicated by an arrow. The position of Daphne Major before 15,000 years ago is shown in red.

The island was chosen for this experiment due to its **Ecological simplicity** - it has low diversity of life-forms, and breeding finch pairs went up to a mere 150

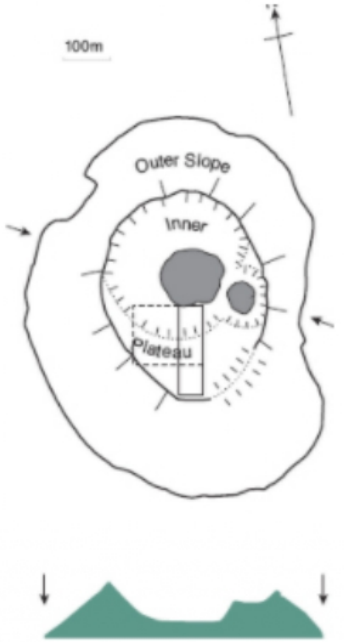


Fig. 2.5 Map of Daphne, drawn from an aerial photograph. The study area used in 1973 for seed sampling, netting, and feeding observations is shown with a dotted-plus-solid line. This was extended in 1977 from the solid line to the outer slope. After 1978 the whole island was treated as the study area except for seed sampling. From Grant and Grant 1980a and Boag and Grant 1984a.

Chp.2. Daphne Finches

Food supply
Name: **Fig. 02-09**
Dataset: **3 attributes, 4 items**
Attribute types: **Categorical, Quantitative**

Fig. 2.9. Uses pie charts to show the seed preferences of two bird species - Fortis and Fuliginosa. Fig. 4.7 shows anatomical drawings of seed shapes and bird beaks. How does the seed shape affect the evolution of bird populations? There is scope for exploded views here.

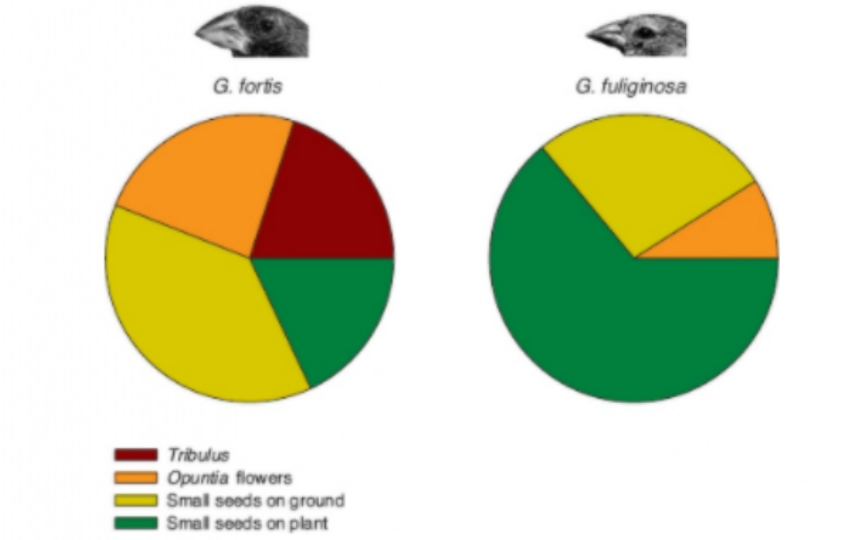


Fig. 2.9 Diets of *fuliginosa* (n = 11 birds) and *fortis* (n = 66) at the end of the dry season on Daphne in 1973. The small seeds on plants comprise *Chamaesyce*, *Tiquilia*, and *Heliotropium* (appendix 1.1). From Boag and Grant 1984a.

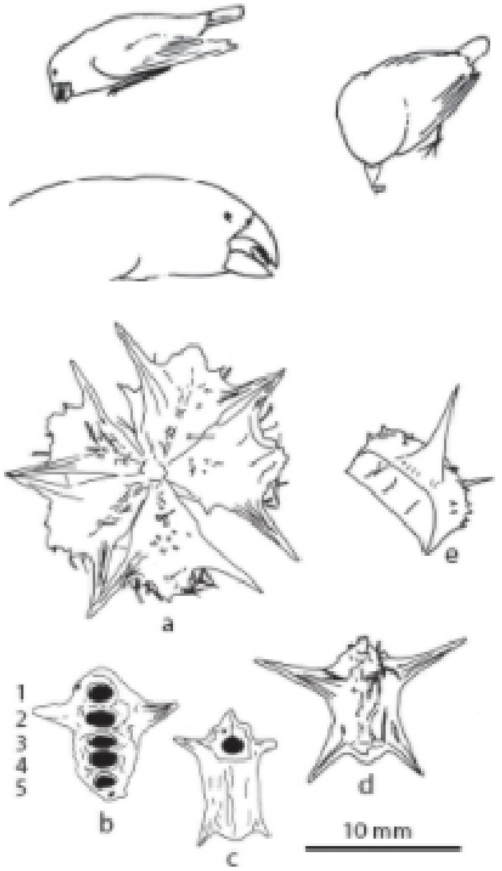


Fig. 4.7 *G. fortis* use different maneuvers to open *Tribulus* mericarps and extract the seeds: either by crushing or by twisting and biting at the corners. Exploited mericarps are shown below. Small individuals sometimes feed parasitically by waiting for a large finch to split a mericarp into two by crushing it, and then seizing an ejected seed. From Grant 1981b.

Chp.2. Daphne Finches

Food supply
Name: **Fig. 02-11**
Dataset: **7 attributes, 10 items**
Attribute types: **Ordinal, Quantitative**

Fig. 2.11. **compares** the 'Cumulative seed abundances' over two islands and three months by **juxtaposing** them. The Y-axis shows the 'Depth x Hardness' classes that the Grants classified based on their calculations. There are 10 classes in all. The depth and hardness of seeds combined with availability or abundance greatly influences the mean beak size of the bird populations.

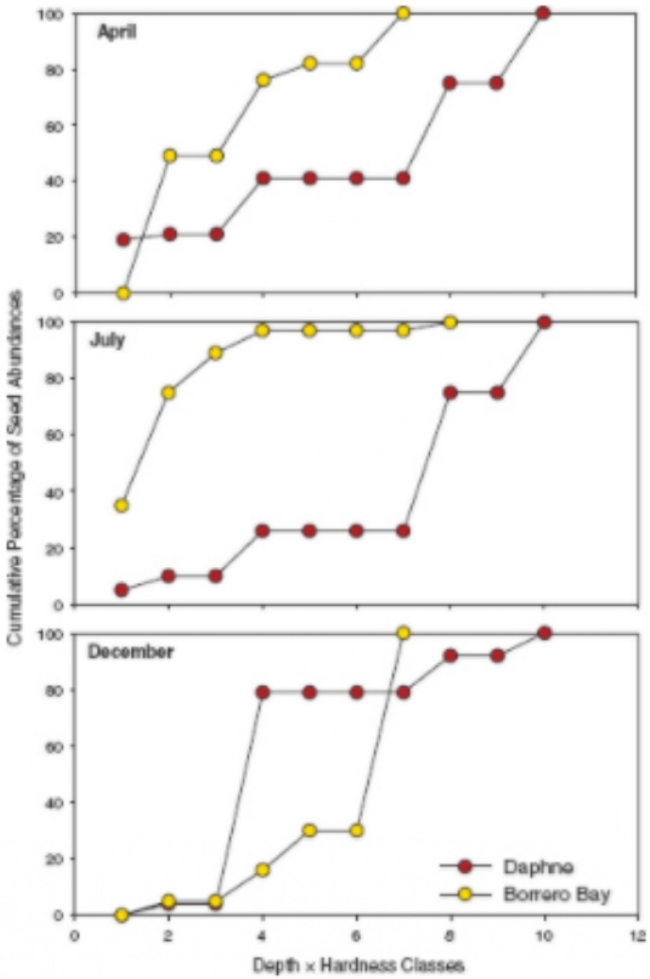


Fig. 2.11 Food distribution profiles in April and December 1973 and July 1975. Hardness classes are the square root of the product of seed depth and hardness (box 2.1). From Boag and Grant 1984a.

Chp.3. Heritable Variation

Name: **Fig. 03-01**
Dataset: **3 attributes, 22 items**
Attribute types: **Ordinal, Quantitative**

This chapter talks about the heritability of certain ‘variations’ in the bird populations. A phenotypic trait is heritable to the degree that the measured variation in the trait reflects variation in genotype.

While I have understood the concept of heritable variation in words, I have yet to understand how this data is calculated and what the terminology means.

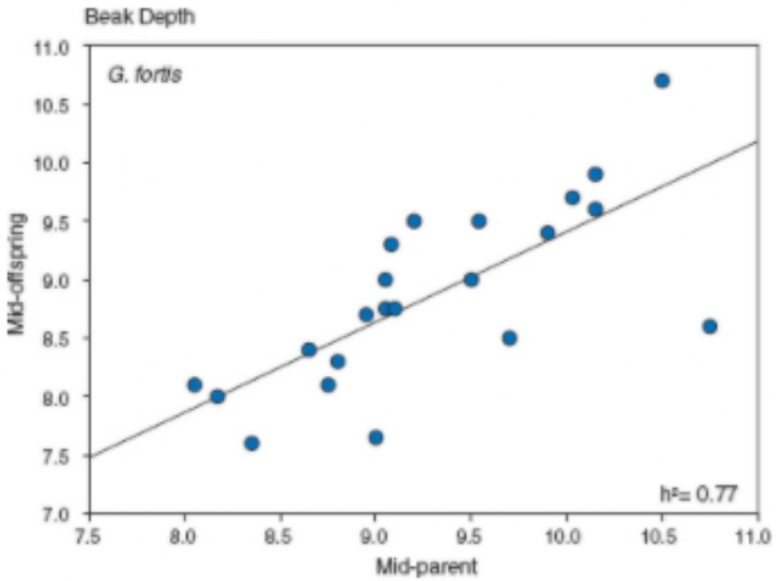


Fig. 3.1 Heritability (h^2) of *fortis* beak depth in 1976, estimated from the slope of the regression of mid-offspring on mid-parent values. From Boag and Grant 1978 and Boag 1983, recalculated with additional measurements.

Family	Mid-offspring	Mid-parents
1	10.7	10.5
2	9.5	9.54
3	9	9.05
4	7.65	9
5	9.4	9.9
6	8.75	9.05

While these are just a few of the datasets and visualizations from the book, it is abundant with more interesting phenomena that have been captured and visualized. A few interesting figures are:

Fig.4.1. The Genotype-phenotype-environment interaction - Concept diagram

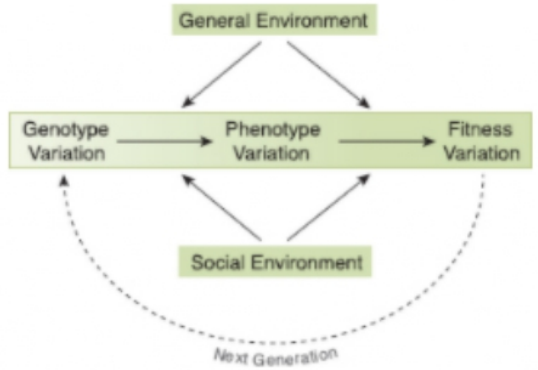


Fig. 4.1 The genotype-phenotype-environment interaction affects fitness. A distinction is made between the social environment of other finches and all other biological and abiotic aspects of the environment. Finches (phenotypes) also affect their environment, for example, when differentially depleting the food supply (this chapter) and destroying the stigmas of cactus flowers (Fig. 2.17; Grant and Grant 1981), and this has feedback effects on fitness.

Fig.4.4. Natural Selection of the fortis body size

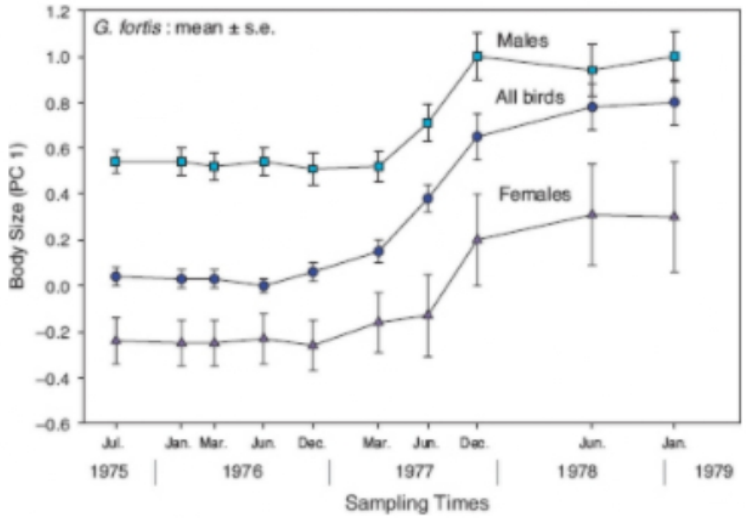


Fig. 4.4 Natural selection on *fortis* body size, indexed by PC1 scores that are explained in the table 3.2 legend. Solid symbols represent means; vertical bars indicate one standard error. Sample sizes of males (n = 198) and females (n = 66) combined with birds of unknown sex (all birds) varied from 642 in June 1976 to 61 in January 1979. There was no breeding in 1977, and the steady increase throughout the drought that year was the result of selective mortality of small birds. From Boag and Grant 1981.

Fig. 4.5. Changes in fortis numbers and its relation to seed abundance and depth x hardness

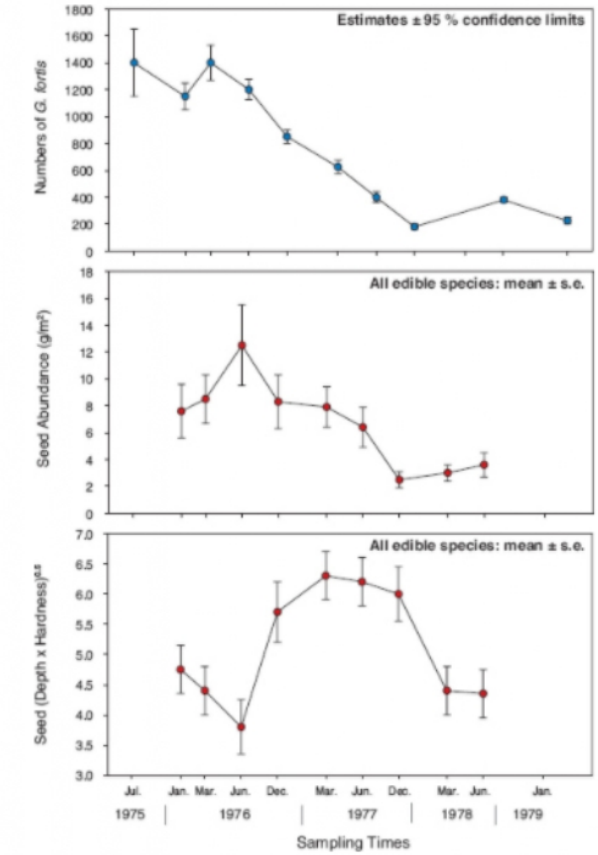


Fig. 4.5 Temporal changes in *fortis* numbers and their principal dry-season food in the drought of 1977. **Upper**: Finch population estimates and 95% confidence limits derived from a Lincoln index based on regular visual censuses of marked birds throughout the island. **Middle**: Estimates of mean seed abundance and one standard error. **Lower**: Estimates of the average size-hardness index ($DH^{3/2}$) of edible seeds and one standard error. *D* is the depth of seed, its second-longest dimension, and *H* is its hardness, measured as the force in newtons needed to crack it open (box 2.1). From Boag and Grant 1981 and Grant 1986.

Fig. 5.2. Changes in average sea temperature, air temperature and rainfall during two months of the El Nino flood.

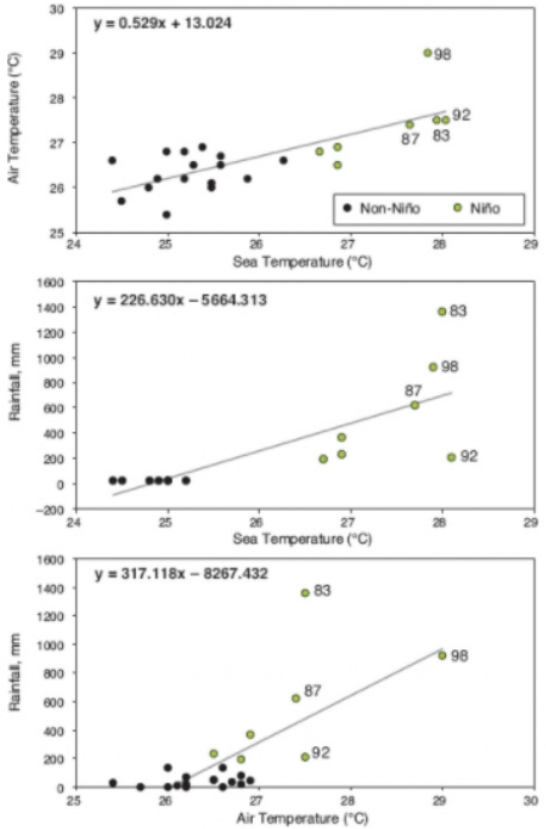


Fig. 5.2 Average daily maximum air temperatures of the warmest two successive months of a year on Santa Cruz Island (1965–98) is a function of the average daily sea temperature of those months at the same location (upper). Annual rainfall on Daphne (1976–98) is a function of the average daily sea temperature (middle) and of the average daily maximum air temperature (lower) recorded on Santa Cruz Island in the warmest two successive months of the year. Seven years of El Niño conditions are shown by green circles. Conditions in four of those years labeled with two digits were clearly extreme, whereas conditions in three others (1991, 1993, and 1997, unlabeled green circles) were not markedly different from the

Fig 5.8 Changes in abundance of caterpillars compared to the numbers of Fortis and Scandens during the El Nino flood.

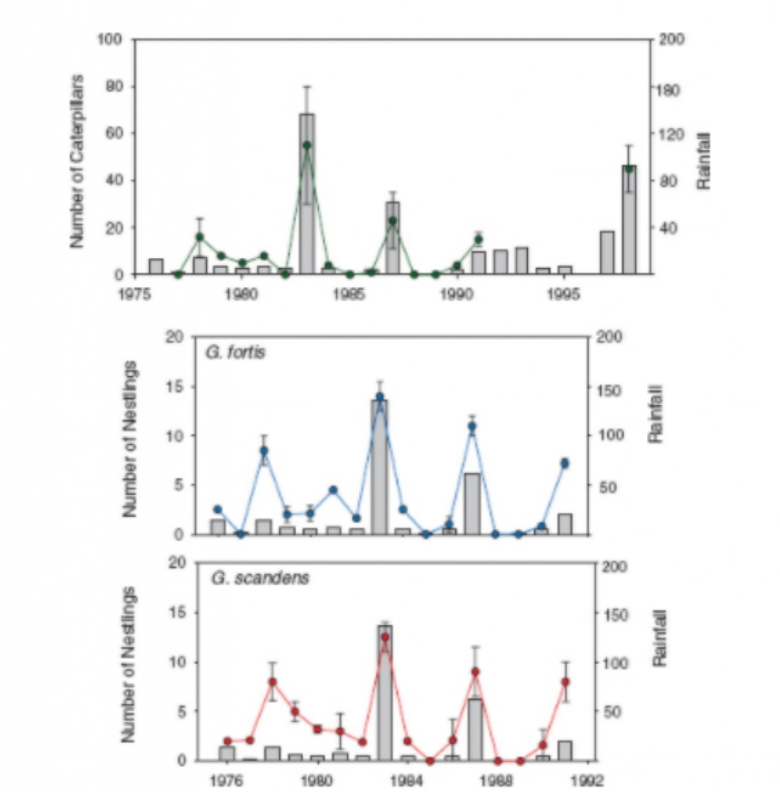


Fig. 5.8 Upper: Annual variation in caterpillar numbers (mean ± s.e. of weekly or biweekly samples) in relation to rainfall shown by histogram bars (from Grant et al. 2000). El Niño years were 1983, 1987, 1991–93, and 1997–98. **Middle and lower**: Annual variation in the mean number of nestlings per female in relation to rainfall. Vertical bars are 95% confidence intervals on the estimates of the means. Only those females breeding in the first month of the breeding season after the first rains are included. Others, usually young, began breeding later in some years, and in a few years there was no breeding at all. Pre-rains breeding by a few *scandens* females have been ignored. For each species mean number of nestlings is highly correlated with rainfall, and so is the mean number of clutches per female (from Grant and Grant 1996b).

Fig. 6.2. The Magnirostris arrives on Daphne Major - Family Tree

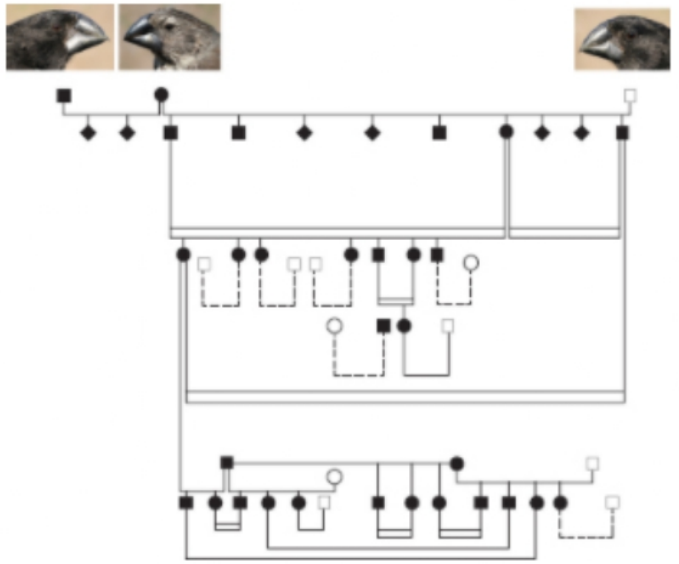


Fig. 6.2 The complete genealogy for *magnirostris* in the years 1983–92. The initial breeders are shown at the top, together with all the fledglings they produced; three of the fledglings became recruits. Below that, only the fledglings that became recruits are shown. Squares, males; circles, females; diamonds, birds of unknown sex. Solid symbols represent banded birds; open symbols, birds without bands. Double horizontal lines indicate breeding pairs of relatives. Broken lines indicate new pairings in 1993 with unknown reproductive outcome. From Grant and Grant 1995b with corrected orientation.

Fig. 6.8. Sonograms of three types of songs sung by the Magnirostris

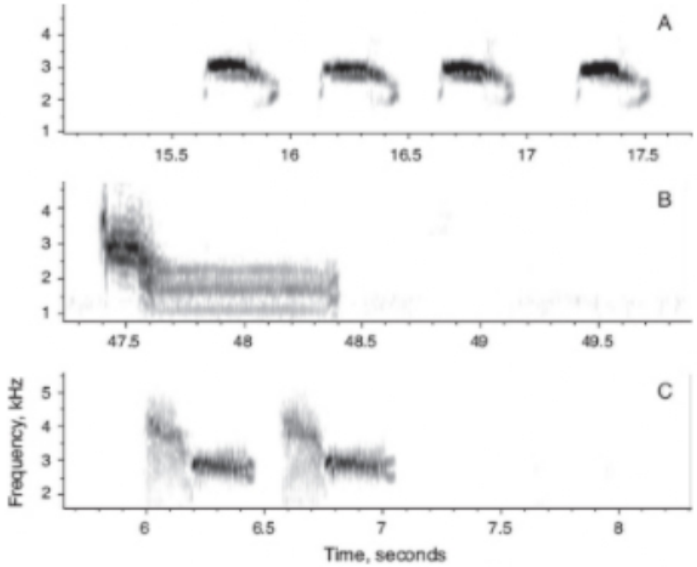


Fig. 6.8 Sonograms of the three types of song sung by *magnirostris* males on Daphne. Note the difference between B and C in which part has maximum amplitude (blackness). From Grant and Grant 1995b.

All the images in this section (Datasets and Existing Visualizations) were taken from the book '40 years of evolution'

Since it is not possible to display all 82 datasets, I have just presented the ones I found most interesting to give a sense of the data available. All the datasets are well-organized and easy to use. Most attributes that require calculation have been calculated beforehand.

Scope for Visualization Design

While the dataset and visualizations provided are quite comprehensive, they leave a lot of scope for better communication and engagement. Some points are outlined below:

Communication

Too dense - For the target audience, these visualizations are low in redundancy. There is a lot of assumed knowledge. The goals of the user would be to get more high-level insights.

Too verbal - There are many verbal descriptions in both the books that can be brought out in the form of concept diagrams and schematic maps. eg: schematic map of the island, measurement process

Not dynamic - There is a lot of scope for using different 'views' to better understand the data. Eg. There is no 'timeline' of events so we are not sure how these evolutionary events interact with each other over time. A timeline view can add a new layer of understanding.

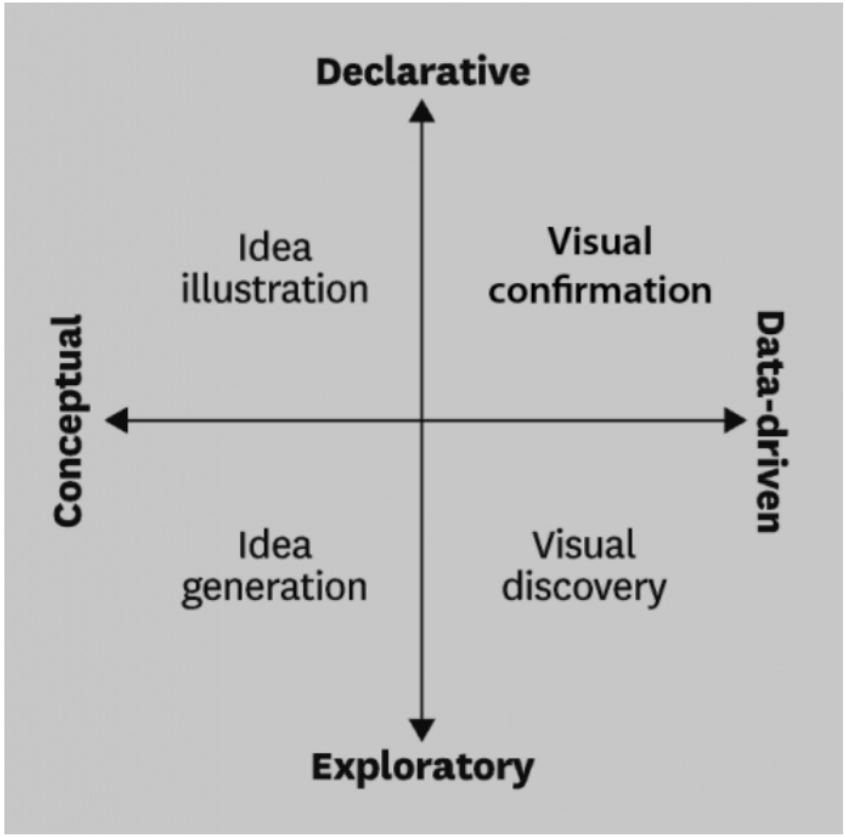
No Layering - While the data provides both high-level insights and specialised insights, there is no layering of information in these visualizations. The information can be layered better to accommodate different knowledge-levels of the viewers.

Engagement

Low in experience - In both the books there are descriptions of the colors, temperature, and weather conditions of the Islands, bleakness of the landscape etc. This sensory information can be brought out better using visual design.

Scope for navigation design - The content is complex enough for a viewer to go back and forth between charts, cross-check, etc. The narrative could be navigated better.

Can be restructured - As a narrative. Instead of a logical sequence of concepts, or what could be considered a top-down approach, the viz can be restructured as a narrative to tell the story of the finches and the researchers over 40 years.



(Image taken from Prof. Venkatesh Rajamanickam's 'Interactive Datavis' course slides)

In order to better plan the visualizations, I looked at 'T**he 4 types of Data Visualization**' quadrants by Scott Berinato. (Berinato, 2016)

Based on the progress so far, the project focuses on **declarative** visualizations, some of which are conceptual and some are data-driven. This would put all visualizations in one of the top 2 quadrants - *Idea Illustration* and *Visual confirmation*.

Methodology and Reading

In this project, the primary devices that I will use to arrive at solutions are ‘The four nested levels of vis design’ and the ‘Three-part analysis framework’ (Munzner, 2015) The four nested levels of vis design outlines the context of the user, i.e., the domain situation, followed by ‘task abstraction’. This will help me understand the ‘goals’ of the user for each visualization. Finally, the third level deals with the visual encoding and interaction idiom used in the visualization to fulfil the goal of the user. The fourth level named ‘algorithm’ will not be relevant in this project since I will not be looking at the computational aspect of the platform.

The three part analysis framework (outlined in the ‘Visualization Design’ section of the report) provides specific keywords and abstractions to better understand and use the process of data and task abstraction and the construction of the idiom.

Apart from these devices, I have also been reading papers on visualization design. Two papers that stood out to me were ‘Reflections on How designers design with data’ (Bigelow et all., 2014) and ‘Interactive Information Graphics’ (Weber, 2017) and I will refer to these throughout the project. The first paper talks about 12 patterns that designers fall into while working with data. The authors make the point that designers tend to follow a visual-first approach rather than a data-first approach. A few reasons for this could be the lack of appropriate tools, lack of familiarity with handling data and a general preference for visual design over data-binding. This paper helped me reevaluate my ideation process. Other resources I am referring to are ‘Envisioning Information’ (Tufte, 1990) and ‘How Charts Lie’ (Cairo, 2019)

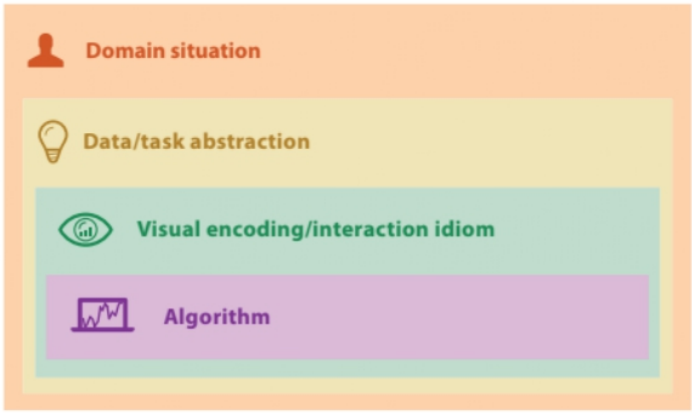


Figure 4.2. The four nested levels of vis design.

Figure 1.7. Three-part analysis framework for a vis instance: *why* is the task being performed, *what* data is shown in the views, and *how* is the vis idiom constructed in terms of design choices.

(above) The images above have been taken from Tamara Munzner’s book ‘Visualization Analysis and Design’

Medium

To narrow down on the medium to be used, I did some secondary research on Data Visualizations for Public understanding of science. Outlined below are the top 4 projects that I looked at in terms of medium.

1. Janet Echelman's 'Earthtime Korea'
Public art installation - experiential
Data-driven
Engagement-driven
Touchpoint: **Public spaces**
www.echelman.com/

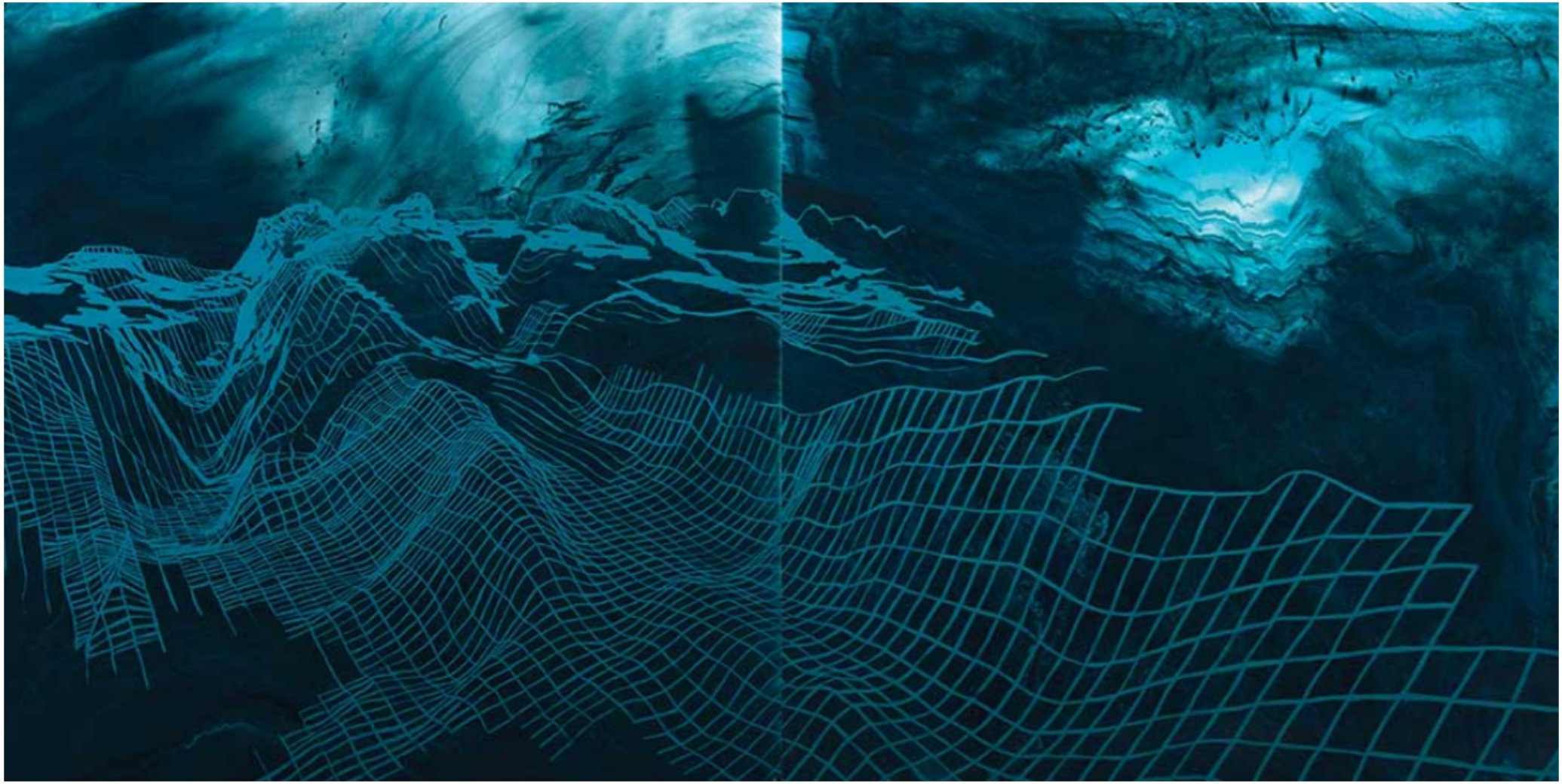
(Echelman, 2020)

"Echelman's studio modeled the physical form for Earthtime Korea using a scientific data set about how a single geologic occurrence in one part of the world had ripple effects all over the world. They measured the change in wave heights of the ocean's surface as they rippled across the entire Pacific Ocean following an earthquake which originated in Chile in 2010. Its form is a manifestation of interconnectedness – when any one element in the sculpture moves, every other element is affected. The earth's day was shortened as a result of this physical event and the length of time measured in microseconds."

(Echelman, 2020)

2. Rebecca Rutstein - Galapagos to San Diego
Data Paintings - experiential
Data-driven
Engagement-driven
Touchpoint: **Gallery**
http://rebeccarutstein.com/

"Rutstein was an Artist in Residence on board the Nautilus science exploration vessel, as it sailed from the Galapagos Islands to Southern California in July 2015.. Rutstein collaborated with scientists who were mapping the ocean floor using multi beam sonar technology, and incorporated the live feed of data into her paintings. Using these data sets, Rutstein has created larger scale works back in her Philadelphia studio."
(Rutstein, 2015)

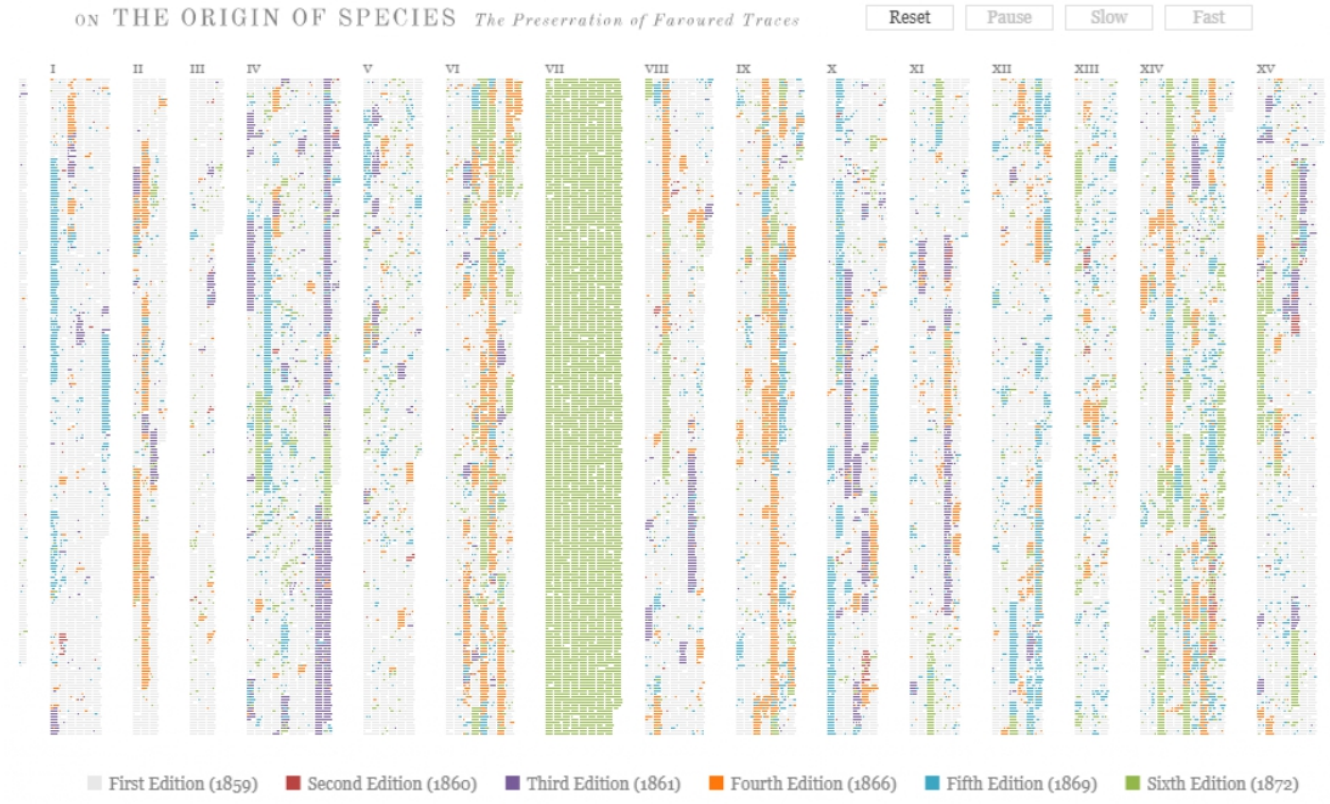


3. Fathom - The preservation of Favored Traces
Interactive Viz + Static Artifact
Data-driven
Communication and engagement
Touchpoints: **Online - phone/PC, bookshops**

https://fathom.info/traces/

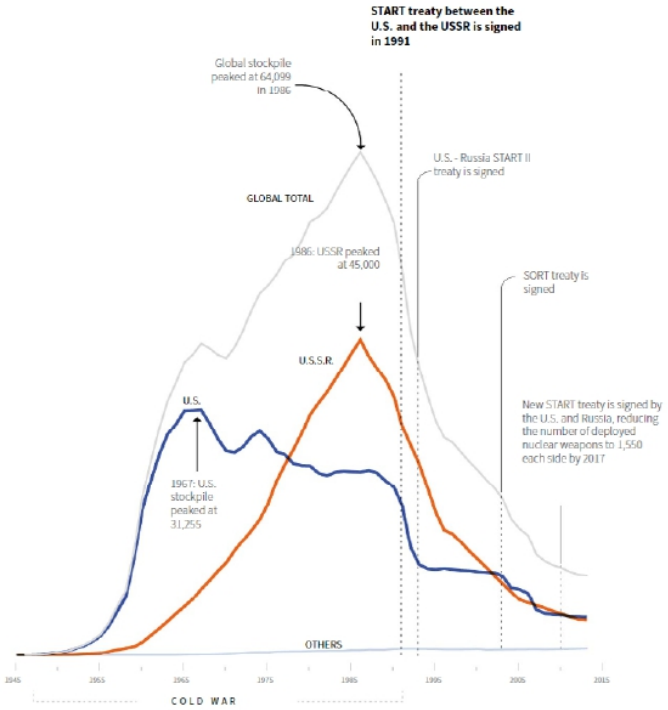
This project also has a similar domain situation to the one I outlined in the 'audience' section of this report.

"By color-coding each word of Darwin's final text by the edition in which it first appeared, our latest book and poster of his work trace his thoughts and revisions, demonstrating how scientific theories undergo adaptation before their widespread acceptance."
(Fathom, 2020)



4. Reuters -In a Flash, a changed world
Data-driven, declarative, static, web-based
Communication and engagement
Touchpoint: **Online - PC**

(Hernandez, 2020)
This Data Story, created on the 75th Anniversary of the Hiroshima bomb blast, outlines the technical details like flight paths, sizes of the bombs, area of damage and so on of the Hiroshima blast.
It also outlines the situation of nuclear weapons today, weapons testing and blast size. What I found interesting was their use of space and scale in the visualizations.



The choice of medium is based mainly on ‘domain situation’ as outlined in the ‘Audience’ section of the report, i.e., to find the right platform to reach the target audience and match their skill-levels in terms of navigation and consumption of the visualization.

Medium: **Web-based**
Devices: **Desktop/Laptop**
Structure: **Narrative - data story**

The next section will be about the ‘Interactivity’ of the Narrative.

Interactivity

Levels of Interactivity

After having chosen the medium, I did some research on the 'Interactivity' of the narrative. In her paper "Interactive information graphics: A framework for classifying a visual genre," Wibke Weber comments on how interpretations of the term 'interactivity' have been 'overworked', and it is more suitable to talk about the different levels of interactivity. (Weber, 2017)

She compiles defintions of the three levels of interactivity in Information Visualizations and their 'Dramaturgic structure.' Rod Sims defines the three levels of interactivity as such:

Low: Object Interactivity
'Object interactivity (proactive inquiry) refers to an application in which objects (buttons, people, things) are activated by using a mouse or other pointing device. When a user "clicks" on the object, there will be some form of audio-visual response.'
(Sims, 1997, p 162).

Medium: Hierarchical and hyperlinked interactivity
'The hierarchical (reactive navigation) class of interactivity can provide the learner with a predefined set of options fromwhich a specific course of study may be selected.'
(Sims 1997,p 163).

High: Immersive virtual interactivity
'provides an interactive environment in which the learner is projected into a complete computer-generated world which responds to individual movement and actions' (Sims 1997, p 167–168).

Why use interactivity?
To decide whether or not to use interactivity, the data's type and denseness come into play. As mentioned in the 'Scope for visualization design' section, the dataset can provide insights for various knowledge-levels from general to specialized. There is scope for layering the visualization insights to cater to the broad audience, which is where a 'medium' level, i.e., hierarchical and hyperlinked interactivity, would be useful.
On the topic of interactivity, Munzner says:
"When datasets are large enough, the limitations of both people and displays preclude just showing everything at once; interaction where user actions cause the view to change is the way forward."

Considerations
While the levels of interactivity are a good starting point, Weber goes on to talk about two more considerations of an interactive infographic:
1. The user's ability to perceive the interactivity - concerned with the user experience design

2. The 'relatedness' of the messages in the interactive sequence (Rafaeli and Sudweeks 1997) - concerned with the dramaturgic structure of the viz.

Dramaturgic structure - Linearity and nonlinearity
Weber explains the term ‘Dramaturgic structure’ as a framework of the infographic that it is closely connected to the ‘Linearity’ or ‘non-linearity’ of the structure. The types are explained below:

Linearity - ‘Linear interactivity . . . refers to applications in which the user is able to move (forwards or backwards) through a predetermined linear sequence of instructional material’
(Sims 1997, p 162–163). It is also described as ‘top-down’ or ‘author-driven’ (Ryan 2006, p 99).

Non-linearity - The meta-message of the nonlinear type is that there is no story and that the users can explore the data or the content by themselves. They can manipulate the graphic by filtering, selecting, and searching the data. (weber, 2017) It is also described as ‘bottom-up-input’ or ‘reader-driven’ (Ryan 2006, p 99).

Hybrid linear-non-linear - The users can choose their own navigation path, but only within a frame predefined by the author or producer. The advantage for the users is that the information is conveyed in a clearly structured way, however with the option to explore the data to a certain degree.
(Weber, 2017)

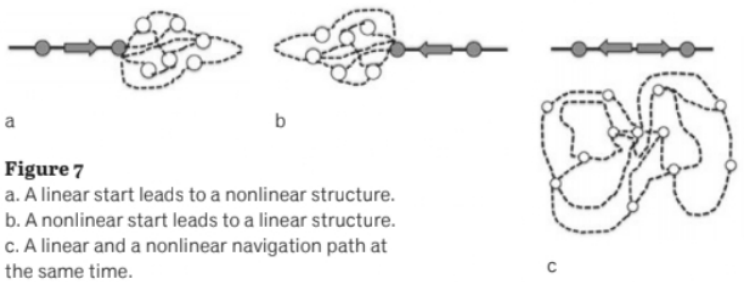


Figure 7
a. A linear start leads to a nonlinear structure.
b. A nonlinear start leads to a linear structure.
c. A linear and a nonlinear navigation path at the same time.

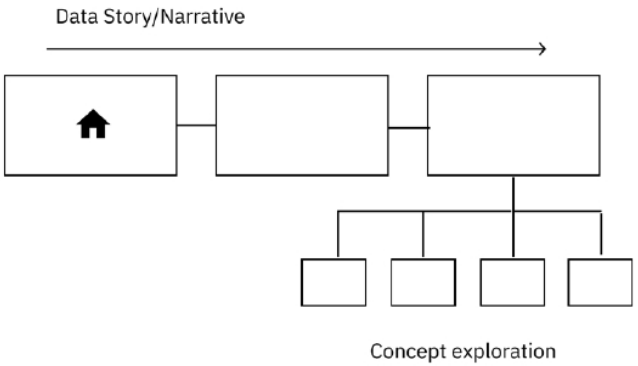
(above) Image taken from Wibke Weber’s “Interactive Information Graphics” (Weber, 2017)

I first ideated four types of Dramaturgic structures (shown on the next page) and then used one of them to make a set of screens to see how it would operate.

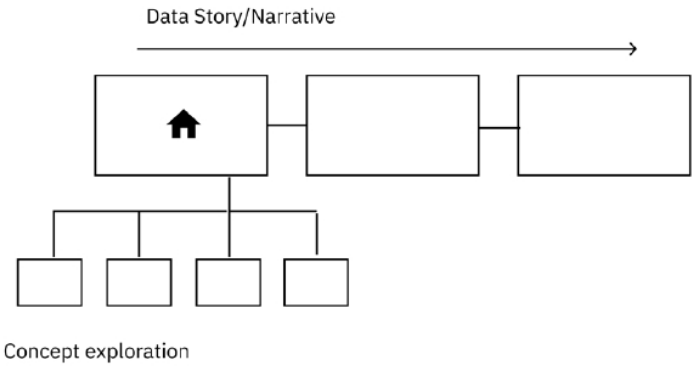
Dramaturgic/Interactive structure explorations

Initial exploration

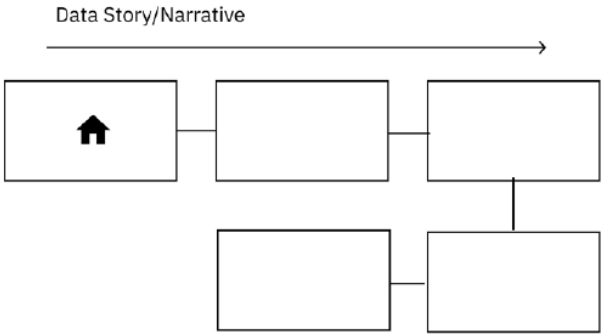
1. Linear > nonlinear hybrid (1)



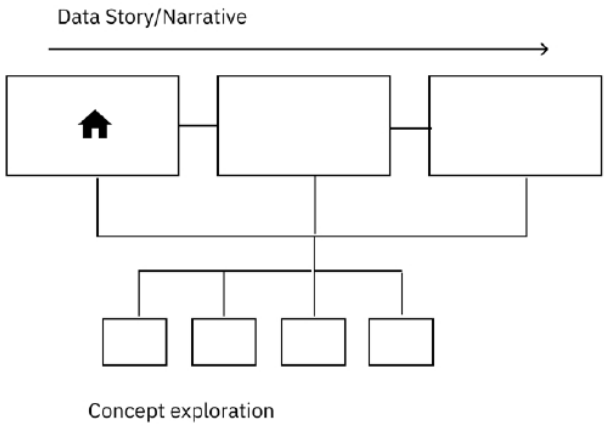
2. Linear nonlinear hybrid (2)



3. Linear



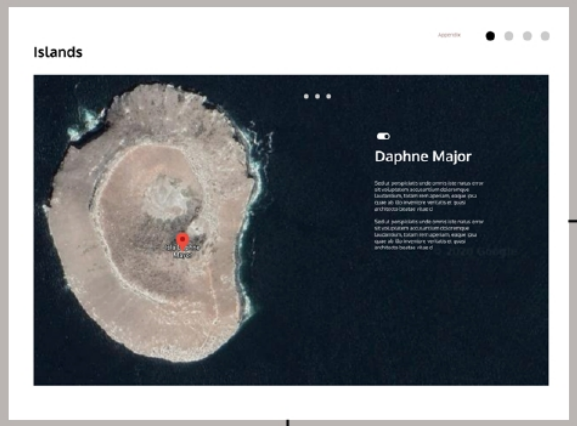
4. Simultaneously linear-nonlinear

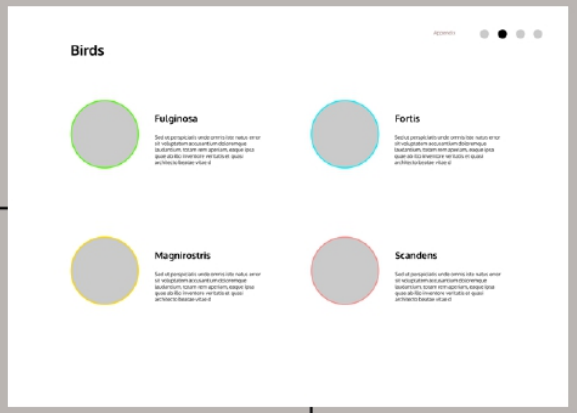


(above) My explorations of Dramaturgic Structures

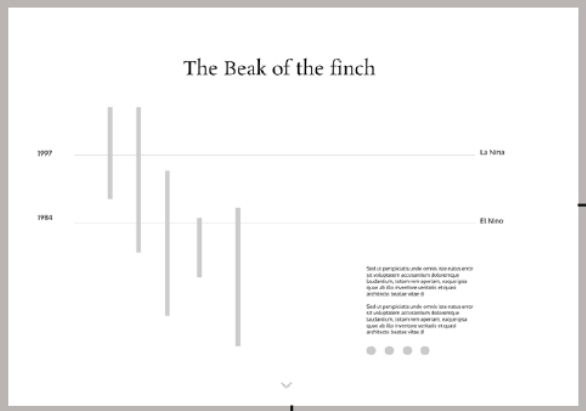
Screens exploring the Dramaturgic structure

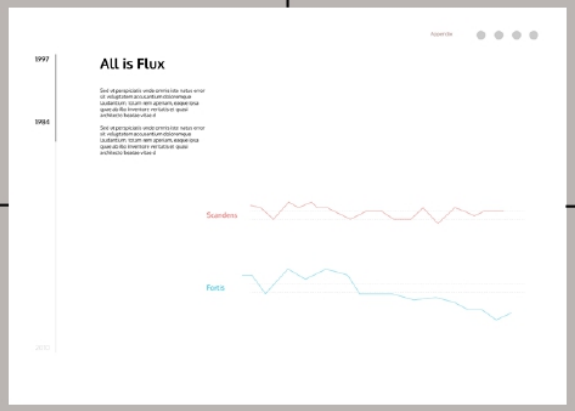
Information cards accessible from any point

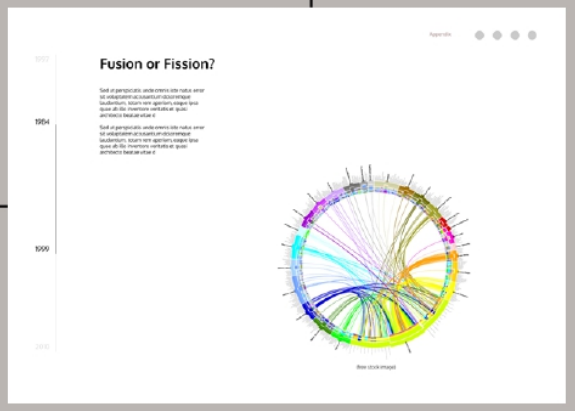




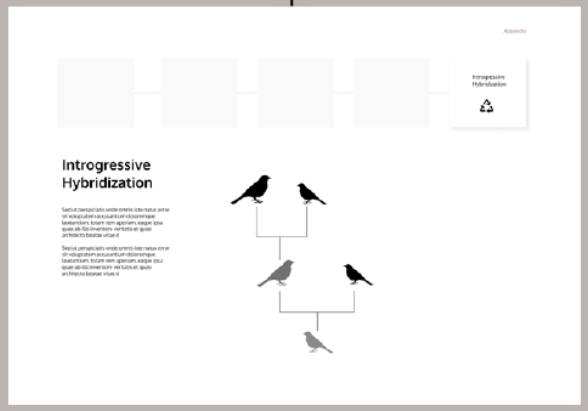
Landing Page







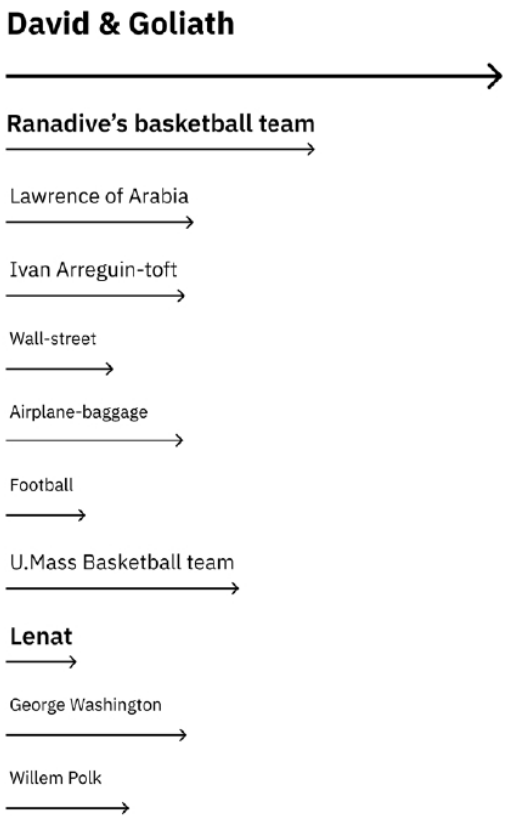
Linear Data Story, Narrative structure



non-linear concept exploration

(above) My explorations of Dramaturgic Structures

Narrative Structure



(above) My take on the "How David beats Goliath" article by Malcolm Gladwell

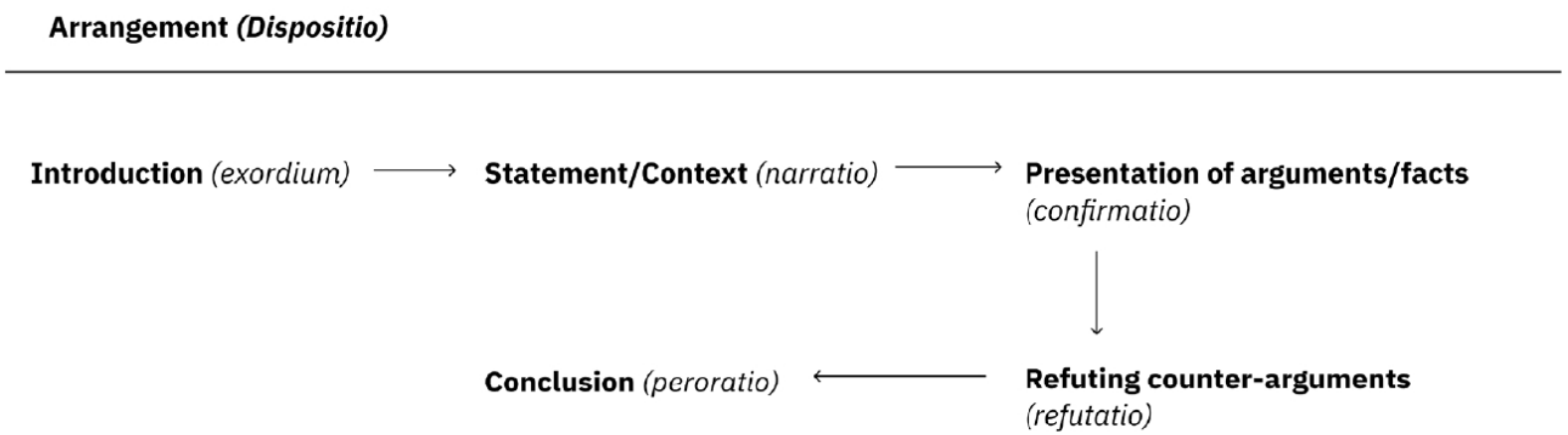
This section is about narrative structure within the dramaturgic structure, i.e. the storyline of the data story. To better understand different narrative structures, I read and took a few notes on narratives that I found to have strong storytelling ability.

1. Malcolm Gladwell
How David beats Goliath
The New Yorker
May 4, 2009

Hook: When underdogs break the rules
Gladwell's article is about how underdogs sometimes win in situations even where the opponent is much stronger.
He makes use of two, even three simultaneous primary storylines. He uses the Biblical metaphor of David and Goliath juxtaposed to Vivek Ranadive's underdog basketball team. He makes frequent diversions to other statistics, and anecdotes that make his point and then returns to the primary storylines.

He ends with "Ranadivé called the press off. He had to. The Redwood City players retreated to their own end, and passively watched as their opponents advanced down the court. They did not run. They paused and deliberated between each possession. They played basketball the way basketball is supposed to be played, and they lost—but not before making Goliath wonder whether he was a giant, after all."

(Gladwell, 2009)



(above) My notes from the "Rhetoric: The Art of Persuasive Writing and Public Speaking" course

2. James Engell
Rhetoric: The Art of Persuasive Writing and Public Speaking
Harvard Edx course

Shown above is the 'arrangement' of arguments in a deliberative rhetoric.

Modes of Appeal
Logos 'Appeal to Reason'- use of deductive reasoning to make a proposition, then look through its logical implications

Pathos 'Emotional appeal' - This is in some sense 'artistic' and relies on the skill of the rhetor.

Ethos 'Ethical appeal' - Establishing the credibility of the speaker or writer. Subtly remind the audience of the rhetor's expertise (could even use specialised terminology)

The course also explained style, schemes, tropes and patterns of sound which I will use in my narrative as well.

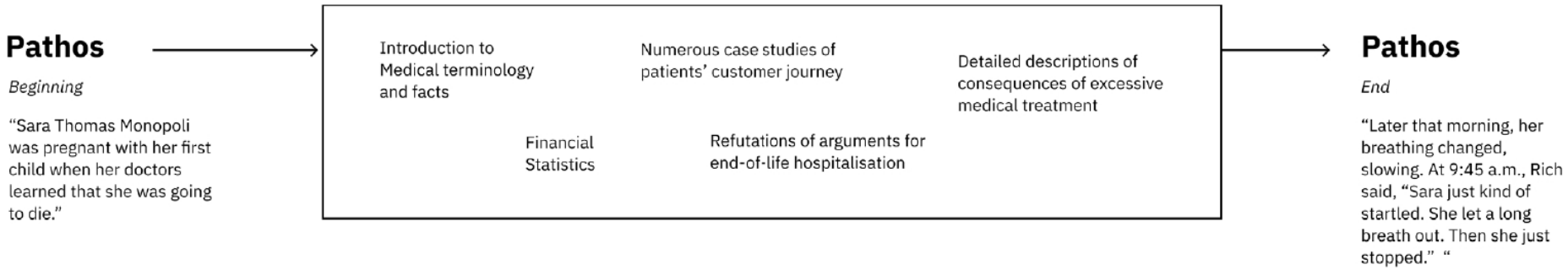
Use of Modes of Appeal

(below) My notes on the use of "Modes of Appeal" in Atul Gawande's article

Letting Go - Atul Gawande, Jul 26, 2010

What should Medicine do when it can't save your life?

"Modern medicine is good at staving off death with aggressive interventions—and bad at knowing when to focus, instead, on improving the days that terminal patients have left "



Examples from the text

Pathos

"More typical was an almost eighty-year-old woman at the end of her life, with irreversible congestive heart failure, who was in the I.C.U. for the second time in three weeks, drugged to oblivion and tubed in most natural orifices and a few artificial ones. "

""I'm running a warehouse for the dying," she said bleakly."

"But doing so represents a struggle—not only against suffering but also against the seemingly unstoppable momentum of medical treatment."

"On the morning of Friday, February 22nd, three days before she was to start her new round of chemo, Rich awoke to find his wife sitting upright beside him, pitched forward on her arms, eyes wide, struggling for air."

Ethos

"I spoke to Dr. Gregory Thompson, a critical-care specialist at Gundersen Lutheran Hospital, while he was on I.C.U. duty one recent evening, and he ran through his list of patients with me."

"But studies suggest otherwise. In one, researchers followed 4,493 Medicare patients with either terminal cancer or congestive heart failure."

"Recently, while seeing a patient in an intensive-care unit at my hospital, I stopped to talk with the critical-care physician on duty, someone I'd known since college. "

Logos

"Medical spending for a breast-cancer survivor, for instance, averaged an estimated fifty-four thousand dollars in 2003, the vast majority of it for the initial diagnostic testing, surgery, and, where necessary, radiation and chemotherapy. For a patient with a fatal version of the disease, though, the cost curve is U-shaped, rising again toward the end—to an average of sixty-three thousand dollars during the last six months of life with an incurable breast cancer."

"But the issue isn't merely a matter of financing. It arises from a still unresolved argument about what the function of medicine really is—what, in other words, we should and should not be paying for doctors to do."

"It's worth pausing to consider what had just happened. Step by step, Sara ended up on a fourth round of chemotherapy, one with a minuscule likelihood of altering the course of her disease and a great likelihood of causing debilitating side effects. An opportunity to prepare for the inevitable was forgone. "

3. Atul Gawande
Letting Go
The New Yorker
Jul 26, 2010

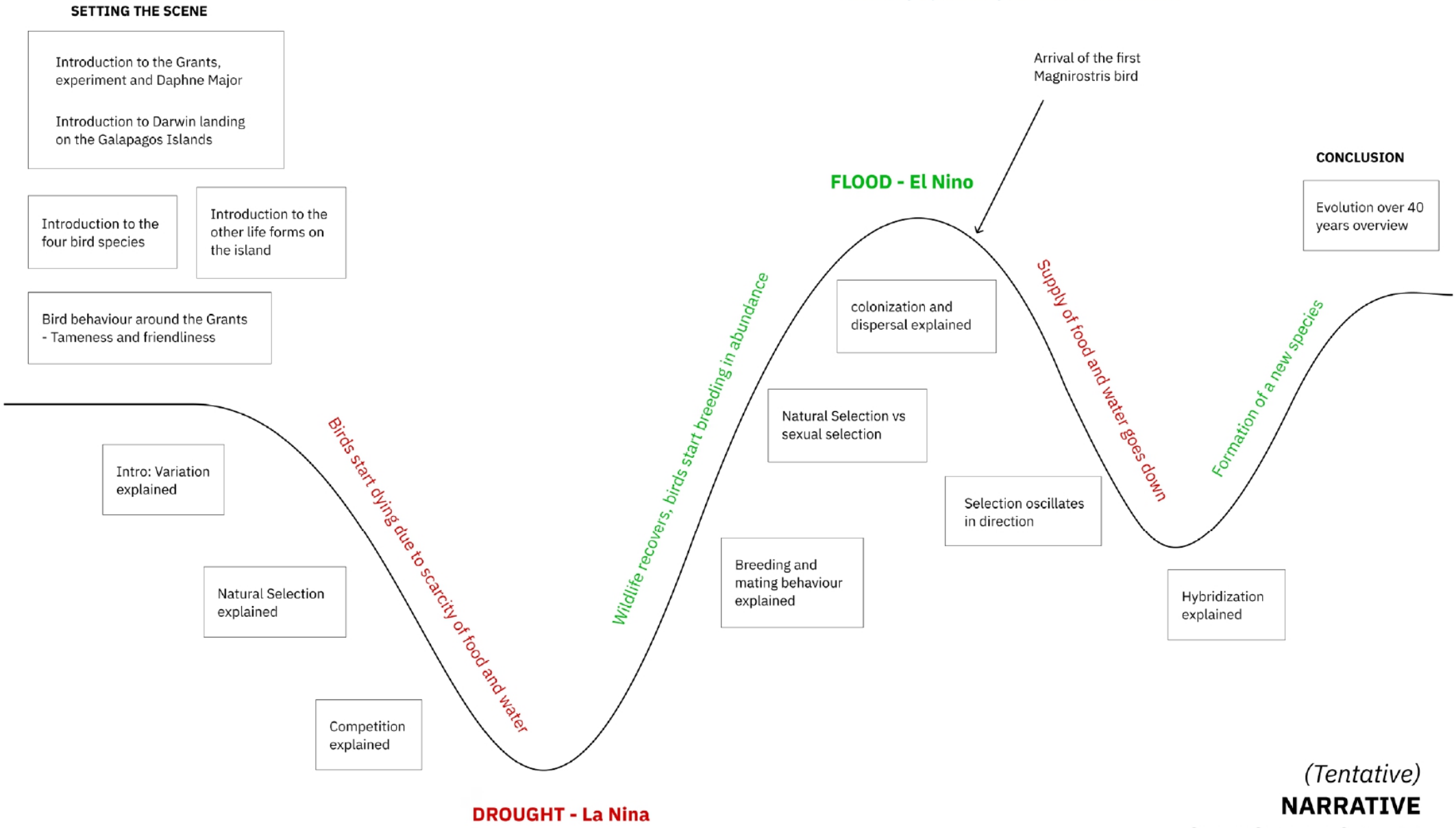
Hypothesis: “Modern medicine is good at staving off death with aggressive interventions—and bad at knowing when to focus, instead, on improving the days that terminal patients have left” (Gawande, 2010)

The article talks about how there is no system in place to stop healthcare for terminally ill patients. There seems to be an assurance that patients can stop treatment whenever they want to, but there is no clarity about when is the right time to stop.

Gawande starts with Sara Monopoli finding out that she is terminally ill and ends the story with her passing away. Throughout this meta narrative, he outlines numerous smaller anecdotes of other terminally ill patients and the ordeals that they had to go through. I also observed that everytime the ‘scientific’ or ‘analytical’ content in his text increases, it is also combined with a lot of ‘pathos’ or ‘emotional appeal’

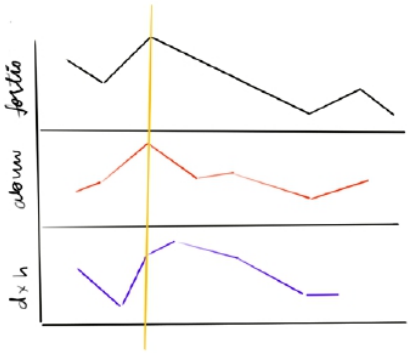
After analysing these narrative structures, I tried to use these learnings to come up with a narrative structure for the Finch story (spread on the next page) The line shows the emotional trajectory of the story. It begins with an introduction to the scientists, birds, islands and life forms on the islands. Then there is an introduction of conflict i.e. the ‘drought’ which causes a lot of birds to die, followed by resolution in the form of good weather. There is conflict again with the dry season, and resolution in the form of the emergence of a new species. Throughout the narrative I have labelled the scientific phenomena that will be explained. This is a tentative narrative.

(below) My exploration of the Finch Narrative structure



(Tentative)
NARRATIVE STRUCTURE OF THE FINCH STORY

Ideation



(above) rough sketches I made while reading the books

I intially started out ideating while I was reading the books. My sketches were purely visual design ideas, and did not follow a data-driven approach. When I read the paper ‘Reflections on how designers design with data’ and Tamara Munzner’s ‘Visualization Analysis and Design’ I decided to change my approach to the ideation process.
I then started to ideating using Munzner’s Three part analysis framework. My approach is to explore the data first by classifying it and then using tools like Tableau to find the shape of the data. After finding the ‘idioms’ I will focus on the Visual design aspect of the visualizations.

One example of the ideation using this process is shown on the following page:

DATA — **TASK** — **IDIOM**

WHAT
Dataset 4.5 -
Data Type: Attributes
Dataset type: Table
Attribute type: Quantitative
Ordering Direction: Sequential
Dataset Availability: Static file

WHY
Actions:
Analyse > Discover
Query > Compare

Targets:
All Data > Trends

HOW (Existing)
Existing Viz:
Facet > Juxtapose

HOW (New)
Existing Viz:
Facet > Superimpose

Existing vis

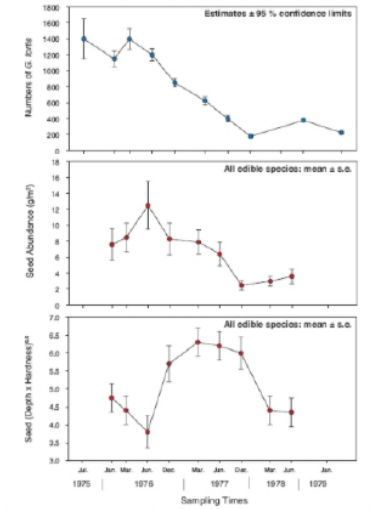
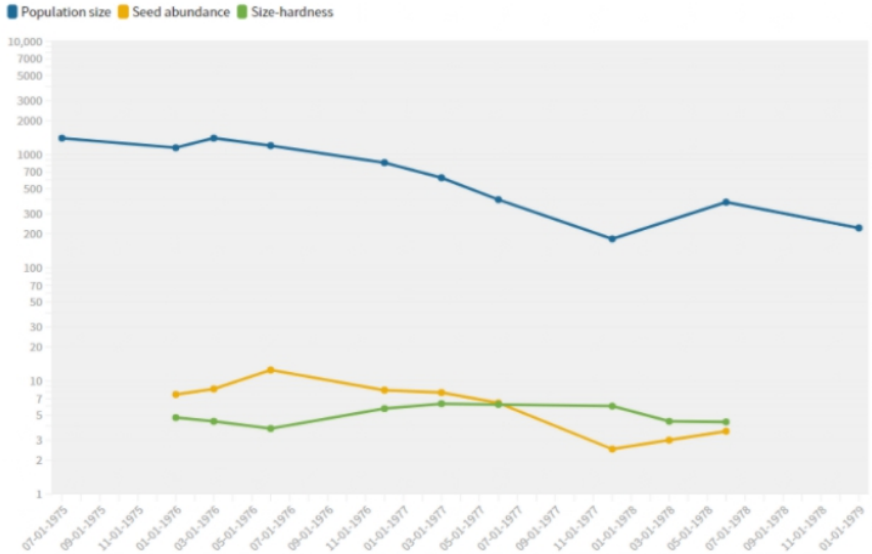


Fig. 4.5 Temporal changes in *fortis* numbers and their principal dry-season food in the drought of 1977. **Upper**: Finch population estimates and 95% confidence limits derived from a Lincoln index based on regular visual censuses of marked birds throughout the Island. **Middle**: Estimates of mean seed abundance and one standard error. **Lower**: Estimates of the average size-hardness index ($DH^{0.4}$) of edible seeds and one standard error. *D* is the depth of seed, its second-longest dimension, and *H* is its hardness, measured as the force in newtons needed to crack it open (box 2.1). From Boag and Grant 1981 and Grant 1986.

New vis



In this case I learned that superimposing all three using a logarithmic scale is not right because the attribute types in all three datasets are different. By superimposing them, I have changed the meaning of the data in a way that can lead to misinterpretations.

Reflections during phase 1

Doing the Data Visualization course, going through readings and ideation so far has led me to reflect on the use of data visualization tools. While I was initially of the opinion that designers only have to focus on the ‘design’ aspects of the visualization process, I have come to realize that a good understanding and exploration of data is required to make any kind of design intervention. To this end, it is necessary for designers to use the right tools to first play with the data, before giving it any kind of shape.

Feedback for Phase 1

During the Stage 1 and Stage 2 juries there were some valid points that were brought up which helped me articulate the project better. I also recieved insightful feedback and ideas from peers, they are mentioned below.

Suggestions

1. Find viewers and informally ask them about their understanding of Evolution, their doubts and questions to get a better idea of the audience's knowledge-level.

2. Look at ways that other people have told scientific stories in an interesting way, eg. Veritaserum or Kurzgesagt. They use a 'hook' or a common myth, and then prove it wrong argumentatively using visuals.

3. Need to explore more examples of dynamic, interesting visualizations. Looking at extreme audiences might be an interesting way to get experimental ideas

4. Articulate whether the contribution of this project is the 'Translation' of the content or whether it is to find new/novel approaches to the data.

Comments

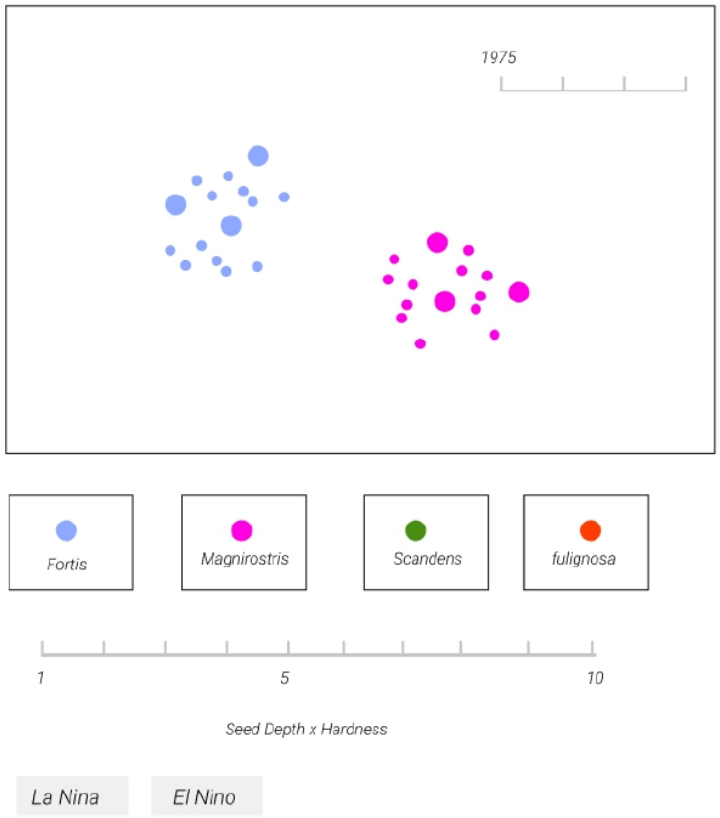
1. The project has good scope for visualization.

2. The explorations done are very interesting, but the planned outcome seems very ordinary and not designerly.

3. It would be better to first come up with the content being used and then decide on the interactive structure.

PHASE 2 - Design and Prototyping

Response to Phase 1 Feedback



(above) DIY 'Evolve your own species' sim ideation

The main feedback from Phase 1 was that the Data Visualizations planned were not new or exciting. The suggestion was to articulate the purpose of the project. Which of these contributions does the project aim to make:

1. Come up with new or Novel visualizations/idioms
2. **Translate specialized content to make it easier to understand for a more general audience.**

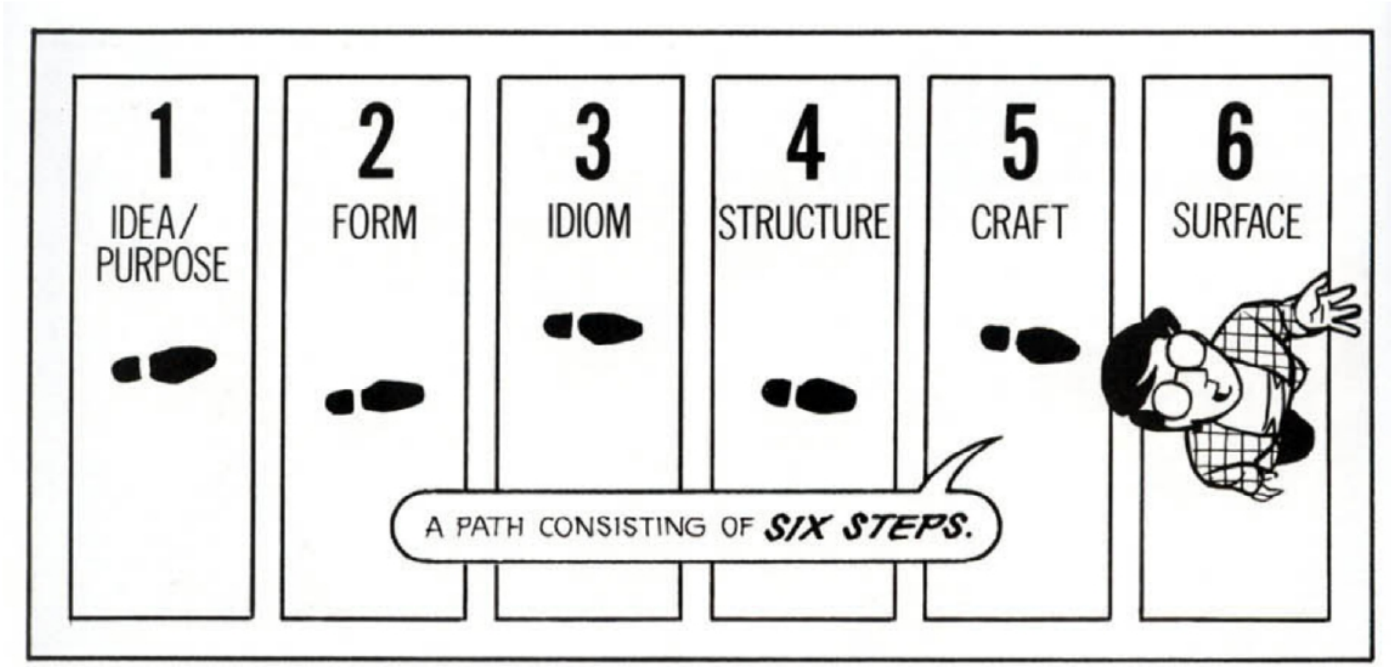
Why not novel visualizations?
Initially, I had explored different ideas of making novel visualizations using the data available. Here are some ideas that I had:

1. Self creating pattern gif similar to Conway's game of life (cellular automata)
2. Data Sound art
3. DIY 'Evolve your own species'

However, the disadvantage of coming up with novel visualizations is that it might compromise on the 'translation' aspect of the content. The reason for picking this particular Evolution study was that it cleared up many misconceptions about Evolution. In keeping with this goal, I decided that the focus of the project is clearly and effectively communicating the content as opposed to having an experimental or novel output.

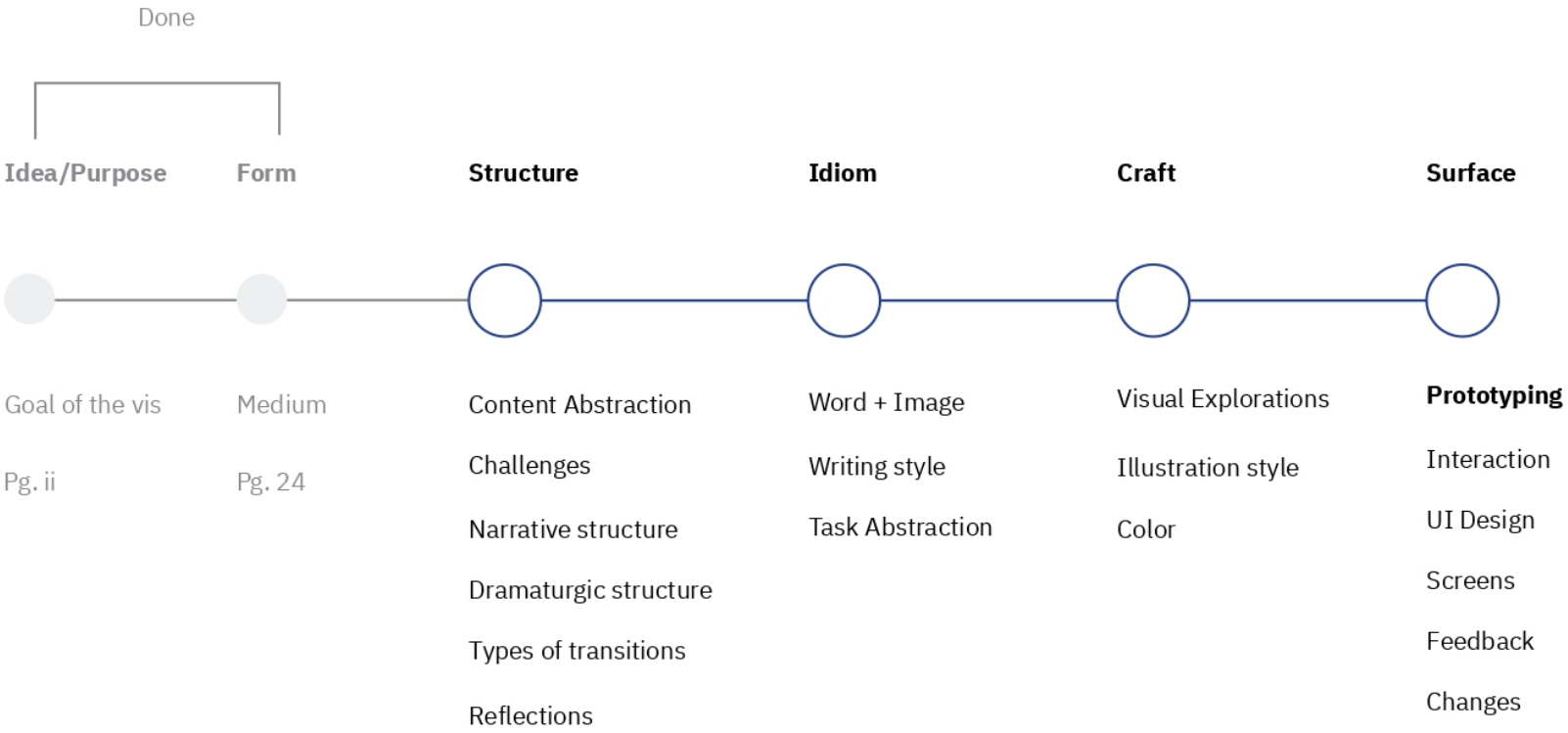
Design Process

Following the explorations of dramaturgic structure, narrative structure and visualizations in Phase 1, I tried to form a Design process which I could follow for the final data story. In the book 'Understanding Comics, Scott McCloud describes comics as "juxtaposed pictorial and other images in deliberate sequence" (McCloud, S., p-9) Since many of his inputs for comic creation applied to data stories, I referred to his design process and reappropriated it to my context while making the final data story.



McCloud, S. (1994). p 170

Design Process



Structure

To structure the data story, it was necessary to first narrow down the content, abstract it so that it may be easy to understand for a lay person, and then work on the narrative structure. Based on the narrative structure, I decided on the dramaturgic (interactive) structure.

Content Abstraction

Since I was working with material from 2 books, the amount of content I went through was vast. *40 years of evolution* was quite technical but provided hard facts and quantitative data. *The Beak of the Finch* on the other hand had high literary quality and was easy to understand but stayed away from statistical information for the most part. My task was to use the narrative from the Beak of the Finch to show the data from 40 years of Evolution.

Challenges

1. *Interdependency of concepts* -
   Scoping down to one is difficult because it depends on prior knowledge, doesn't give the full picture. Eg. Fuliginosa went extinct 15,000 years ago, therefore today, the Fortis size is smaller.

2. *Simultaneous events* -
   For any evolutionary outcome to be formed, there are multiple forces simultaneously at play. How can that be expressed in the narrative without complicating the explanation?

3. *Preserving details/complexity* -
   While generalizing the content, to get high level insights, there is a risk of oversimplification. To avoid misrepresenting it as a simple straightforward process, it is necessary to preserve the details.

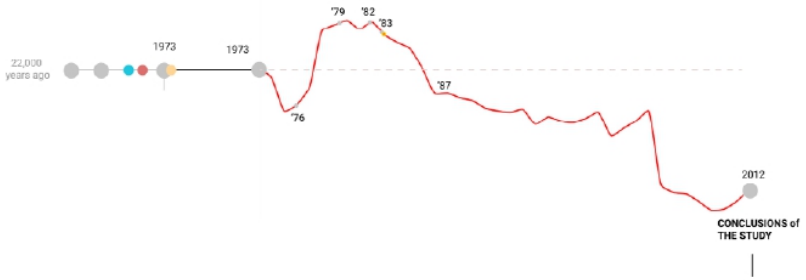
4. *Levels of insight* -
   As viewers have different knowledge elevls, they might need to view the information at different levels of complexity. How can the content be abstracted for such a use case?

These content abstraction challenges were overcome by designing a timeline-based narrative.

Narrative Structure

The narrative structure outlined in Phase 1 provided a good starting point for the data story, however, it did not account for the interdependency of concepts in the story. To solve these challenges, I found that a 'Timeline' of events unfolding starting from 20,000 years ago, and ending in the present would be more appropriate.

(continued >)



(above) Navigable timeline ideation

Timeline

1. *Makes interdependency visible* -
   The timeline acts as 'notes' for previous events which can be recalled as the story progresses. This makes it easy to refer back to a previous event.
   eg. Fulignosa extinction and character release

2. *Makes simultaneity visible*
   Throughout the study, many phenomena occur simultaneously.The timeline makes it easy to represent this in a linear narrative without losing track of the story.
   eg. Magnirostris starts breeding in 1983, the same year that the Fortis see a sharp increase in population.

3. *Preserving details/complexity* -
   In some cases, the scientific content is too dense and it is better not to elaborate on it in the narrative. But the fact that it did occur should be mentioned to bring out the complexity of the process. The timeline allowed me to mention events in passing without elaborating upon them too much. Eg. Separation of Daphne major from the main island.

4. *Levels of insight* -
   Timeline view gives an overview of the process, while the description below elaborates on the events. This allows viewers to choose the granularity at which they want to see the information.

5. *Progressive disclosure*
   Because the process of evolution can be complex and multilayered, the timeline-based narrative discloses information in smaller chunks, from event to event so as not to overwhelm the viewer.

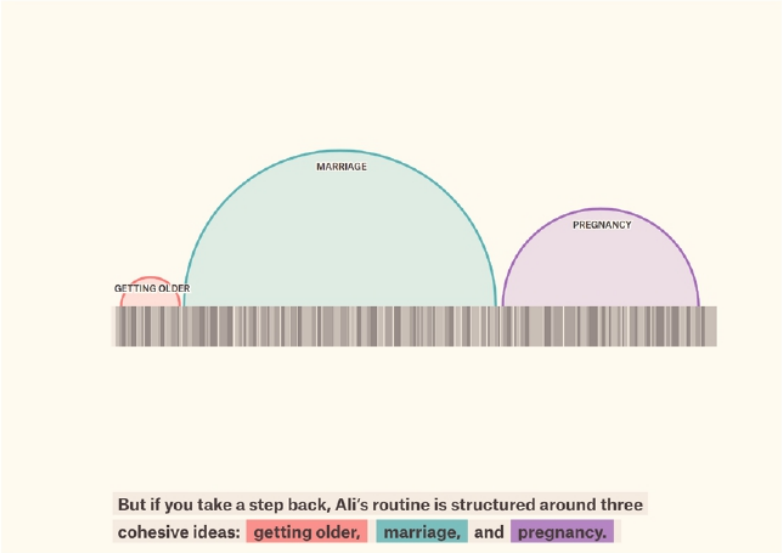
6. *Medium as the message*
   The study illustrates that Evolution is not an abstract concept from millions of years ago but is happening presently in our very backyards. The timeline takes the viewer from 20,000 years ago till the present and illustrates that Evolution is in fact more visible in

As it stands the narrative begins with Darwin's idea of Evolution, and then introduces the reader to the Grants, Daphne major island and the Finch birds along with the concept of speciation. Then the timeline kicks in and carries the narrative from 20,000 years ago up till 2012. The narrative closes with questions of the relevance of Evolution today in 2020.

Dramaturgic structure

In phase 1, I had explored different ‘linear’ and ‘non-linear’ types of interactive/dramaturgic structures. Based on the narrative structure, I found it best to stick to a linear dramaturgic structure so as not to confuse the viewer. While designing it I also realised that it is quite difficult to switch from a linear to non-linear narrative and that is beyond the scope of this P2.

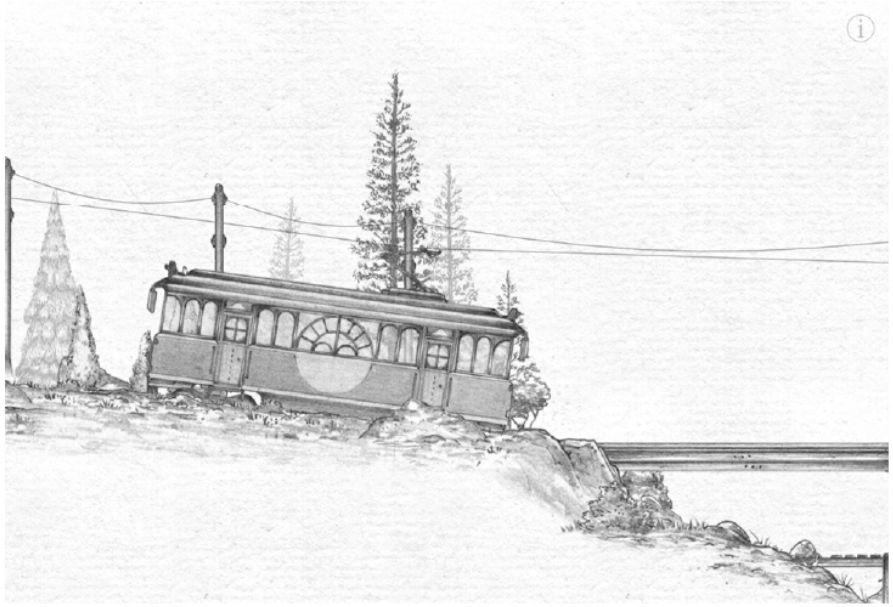
(below) A few explorations suggested by Prof. Jayesh that inspired me to go for a timeline navigation and linear dramaturgic structure.



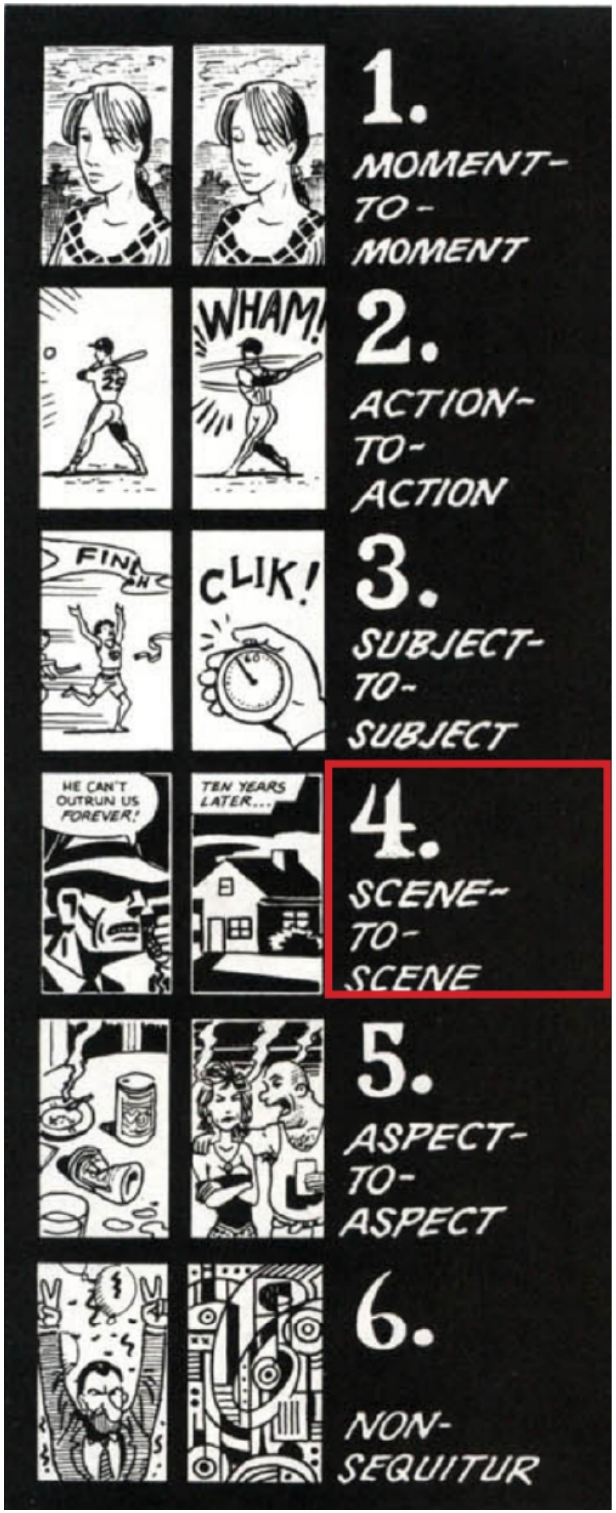
Pudding.cool

‘The Structure of Stand-up comedy’

by Russell Goldenberg & Matt Daniels



‘Short trip’ by Alexander Perrin



McCloud, S. (1994).
p 74

Types of Transitions

Shown on the left are different transitions from the book 'Understanding comics'. McCloud talks about the pacing of the comic and the passage of time captured between the frames, throughout the narrative. I found this relevant to the Data Story as I also needed to design the pacing of the story. The transition that comes closest is 'Scene-to-scene' however, I would describe it more precisely as 'event-to-event'.

Reflections

While coming to some of these decisions I realised that the content abstraction, narrative design etc. are decisions that inform each other and happen simultaneously. In my case, desigining a narrative was an intuitive process, and not preplanned. The decisions can most easily be analysed and justified once the narrative has already been constructed.

Idiom

Both Munzner and McCloud use the word ‘idiom’ to refer to a mode of expressing the data or visual. Designing the ‘idiom’, as I understood it would mean thinking about the design of individual visualizations, writing style and the relationship between words and visuals.

Writing

These are a few writing goals I had:

1. *Convert heavy vocabulary into easy-to-understand sentences*
   Reduce complexity, offload the weight of heavy vocabulary into smaller simpler words.
   Eg: Sentence structure : Reproductive preferences > how they choose to reproduce.

2. *Make the text sound approachable*
   eg. His explanation of speciation laid stress on three factors > He explains that there are 3 main reasons for species to form.

3. *Avoid Terminology unless necessary* -
   Eg: Heritable variation > They found that some characteristics could be inherited more easily than others.

Writing Challenges

1. Risk of scientific inaccuracy while simplifying words
2. Maintaining the balance between too dense and too wordy.
3. During Phase 1 I got feedback that introducing a fictional character or storyline might help with the storytelling. However, I thought that it might lead to misconceptions or blur the line between fact and fiction. For those reasons I did not attempt fiction writing.

Visualization - Task Abstraction

Earlier in the report I had provided an example of Munzners Three-part analysis framework and how it can be used to come up with a visualization idiom. (Pg. 39) I followed the same framework while ideating for the idioms. The next few pages illustrate them.

Beak width and length for 4 bird species, across islands

Data: Table with Island-wise averages of Beak width and length

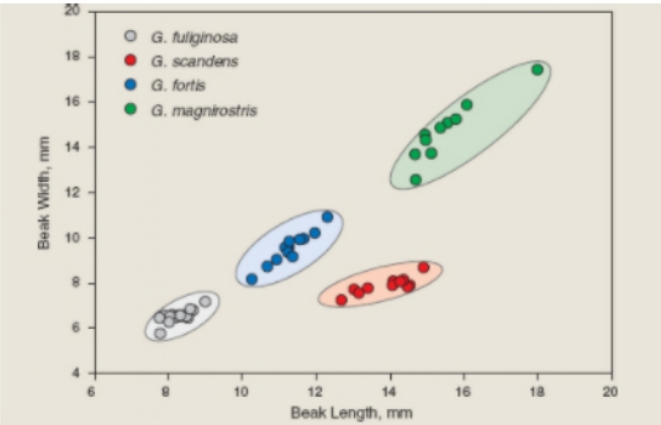
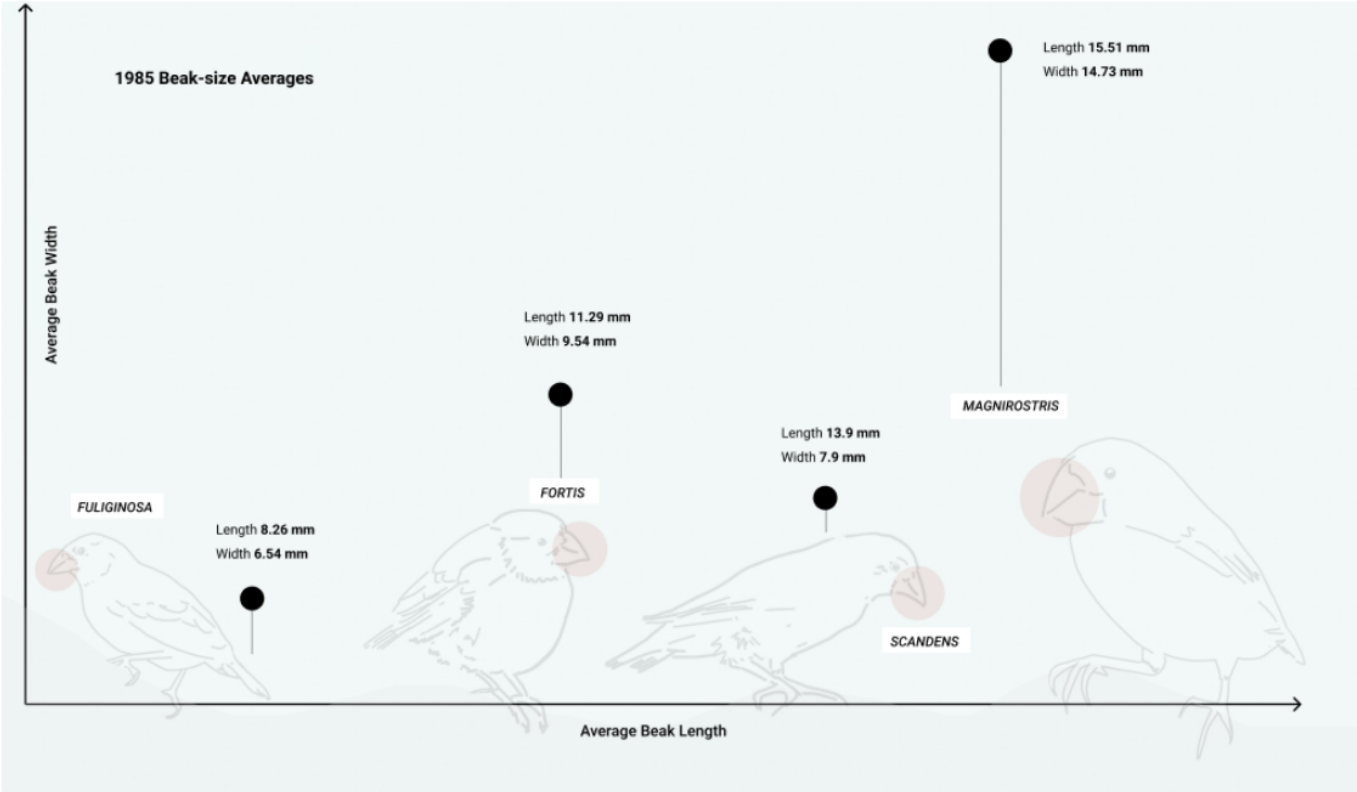


Fig. B.1.2 Morphological variation among four species of Darwin's ground finches (males) on several islands. Data are taken from Grant et al. 1985.



OLD VIS

Old Viz: Density Ellipses plotting averages across islands. This viz was meant for science students/ practitioners to understand the distribution of bird beak within and between species.

NEW VIS

New Viz: Scatter plot showing the overall average beak size for all four species. Meant for viewers who are only concerned with betwen-species beak sizes comparison.

P.4 Phenotypic variation in the *G. fortis* population on Daphne.

OLD VIS

Old Viz: Goal is to bring out the differences in phenotype within a species. Birds are arranged in no particular order.

Phenotypic variation in Fortis +
Phenotypic variation in Scandens

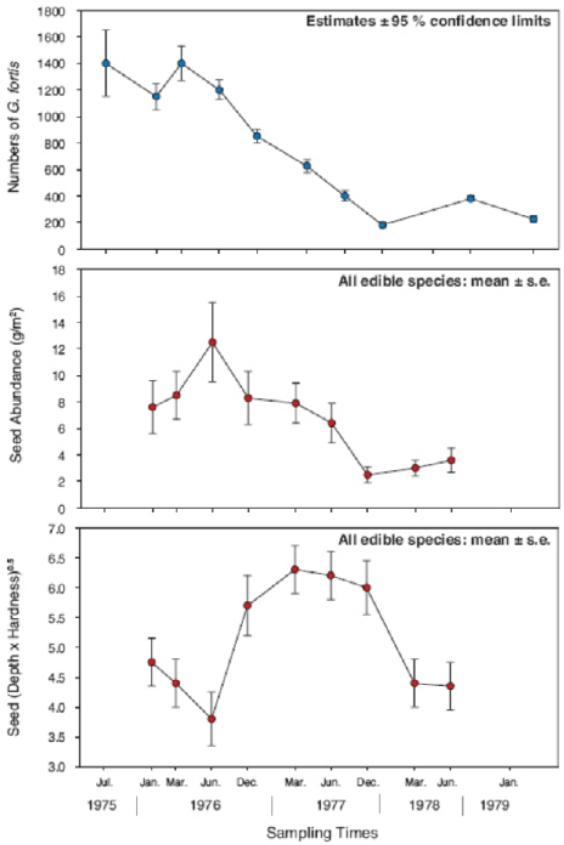
Data: Photos of different birds in two species

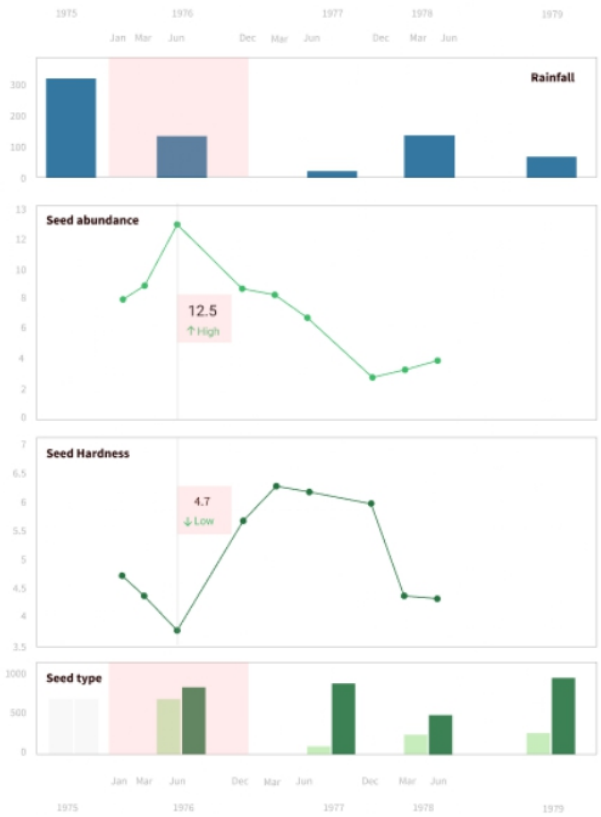
NEW VIS

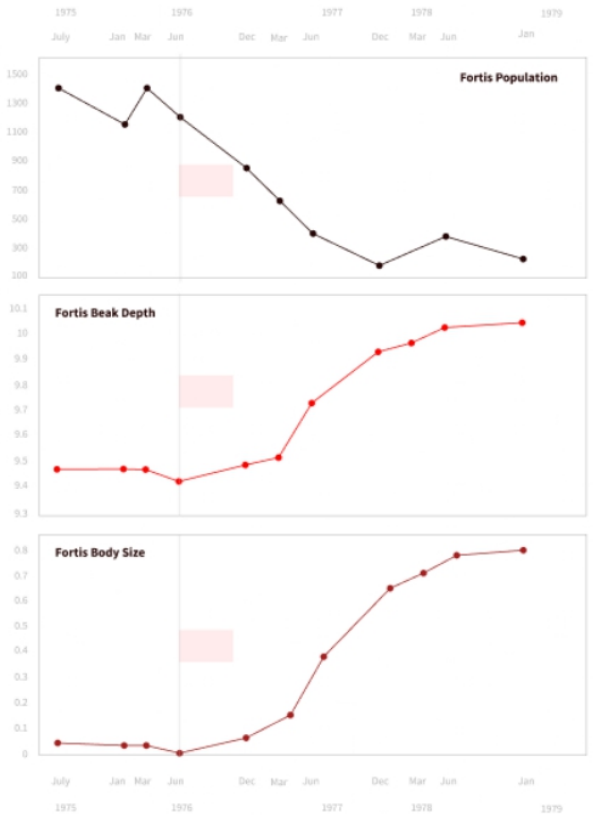
New Viz: The goal is to notice the phenotypic variation within species, notice the difference in beak shape between species and also show the beak size range within species. The birds are visually sorted by beak size. Their beaks poking out of the frame emphasize the beak shape and size.

Correlation between Population, seed-abundance and seed hardness.

Data: Quantitative







OLD VIS

Old Viz: Goal is to show the correlation between the three charts.

NEW VIS

New Viz: The goal is understand the effect of rainfall > seed abundance > seed hardness > Fortis population > beak size > body size.
Viz allows the user to hover over the chart and track the entire journey.

Genealogy for Magnirostris between 1982-92

Data: Tree

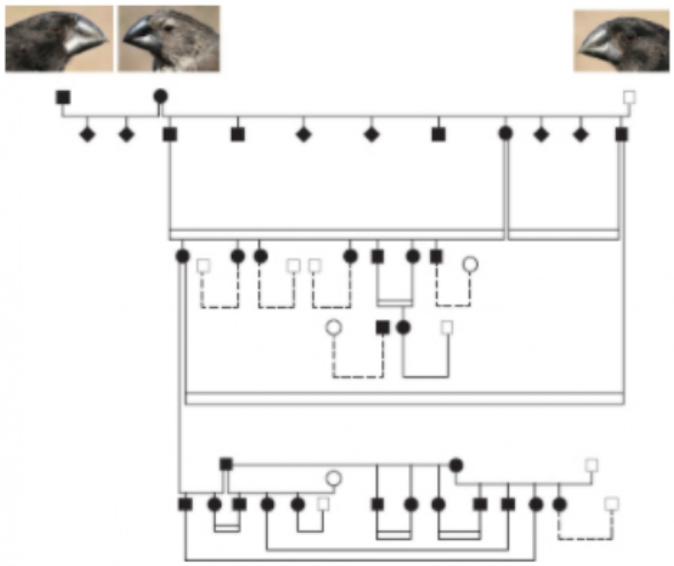
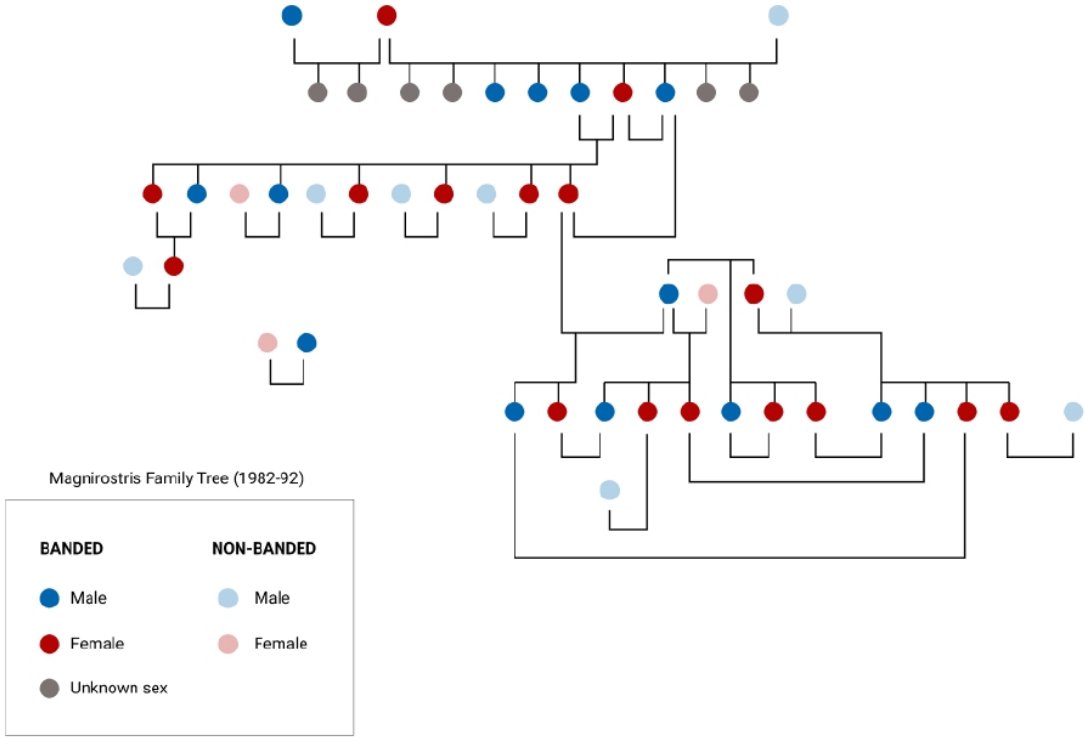


Fig. 6.2 The complete genealogy for *magnirostris* in the years 1983–92. The initial breeders are shown at the top, together with all the fledglings they produced; three of the fledglings became recruits. Below that, only the fledglings that became recruits are shown. Squares, males; circles, females; diamonds, birds of unknown sex. Solid symbols represent banded birds; open symbols, birds without bands. Double horizontal lines indicate breeding pairs of relatives. Broken lines indicate new pairings in 1993 with unknown reproductive outcome. From Grant and Grant 1995b with corrected orientation.

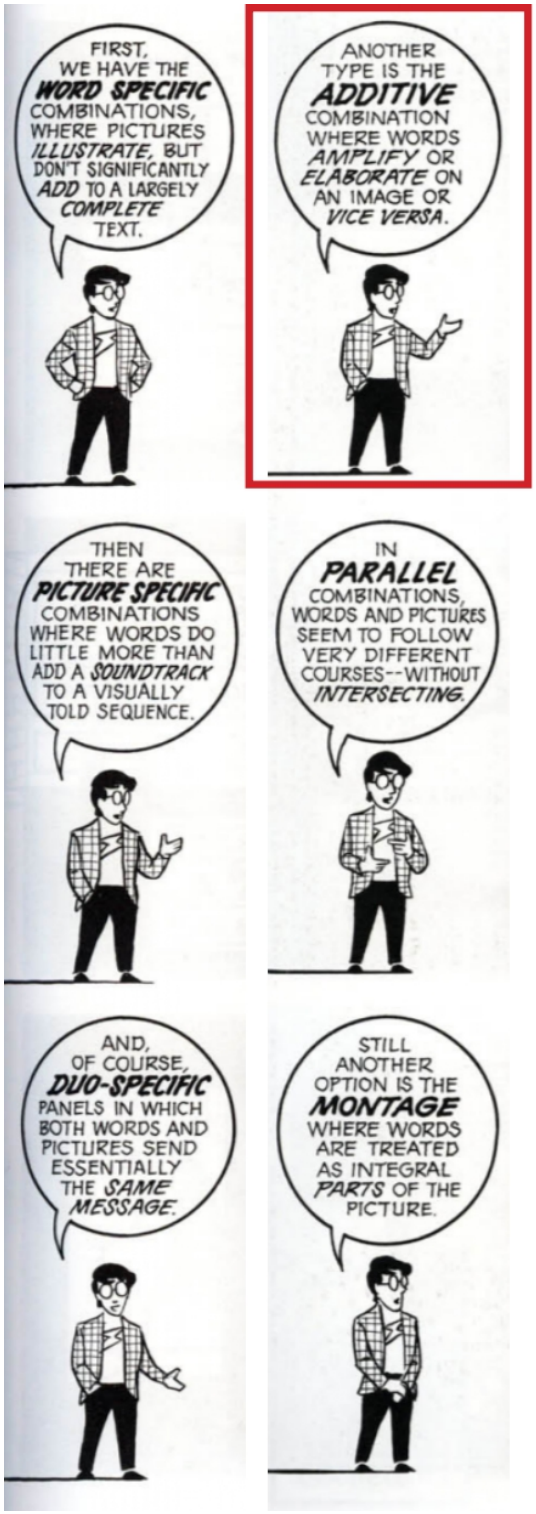


OLD VIS

Old Viz: Goal is to show the growth of the population, spread of genes, and inbreeding. Encoding for gender is shape, encoding for 'banded' is 'fill'. Spacing/layout is not efficient in bringing out the relations.

NEW VIS

New Viz: The goal is to see the population growing, and explore/have fun with the following trivia - how many males, how many females were present, which birds are related, etc. The encoding for gender is color. The encoding for banded/non-banded is opacity.



McCloud, S. (1994).

p 153

Word + Image

McCloud talks about the relationship between visuals and image in the narrative. While one is not better than the other, I thought it was important to articulate and maintain one standard combination throughout the story. In this data story, the combination that makes most sense is the Additive type where the visual takes precedence and the text elaborates or amplifies what is being shown.

Craft

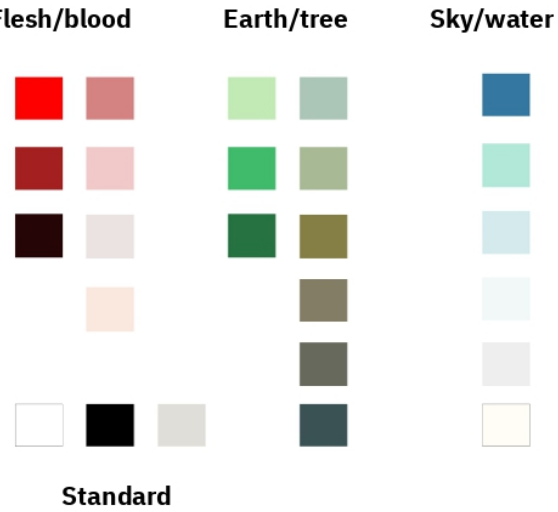
In my understanding ‘Craft’ refers to the use of visual design principles, color, treatment, etc. I looked at three aspects of craft:

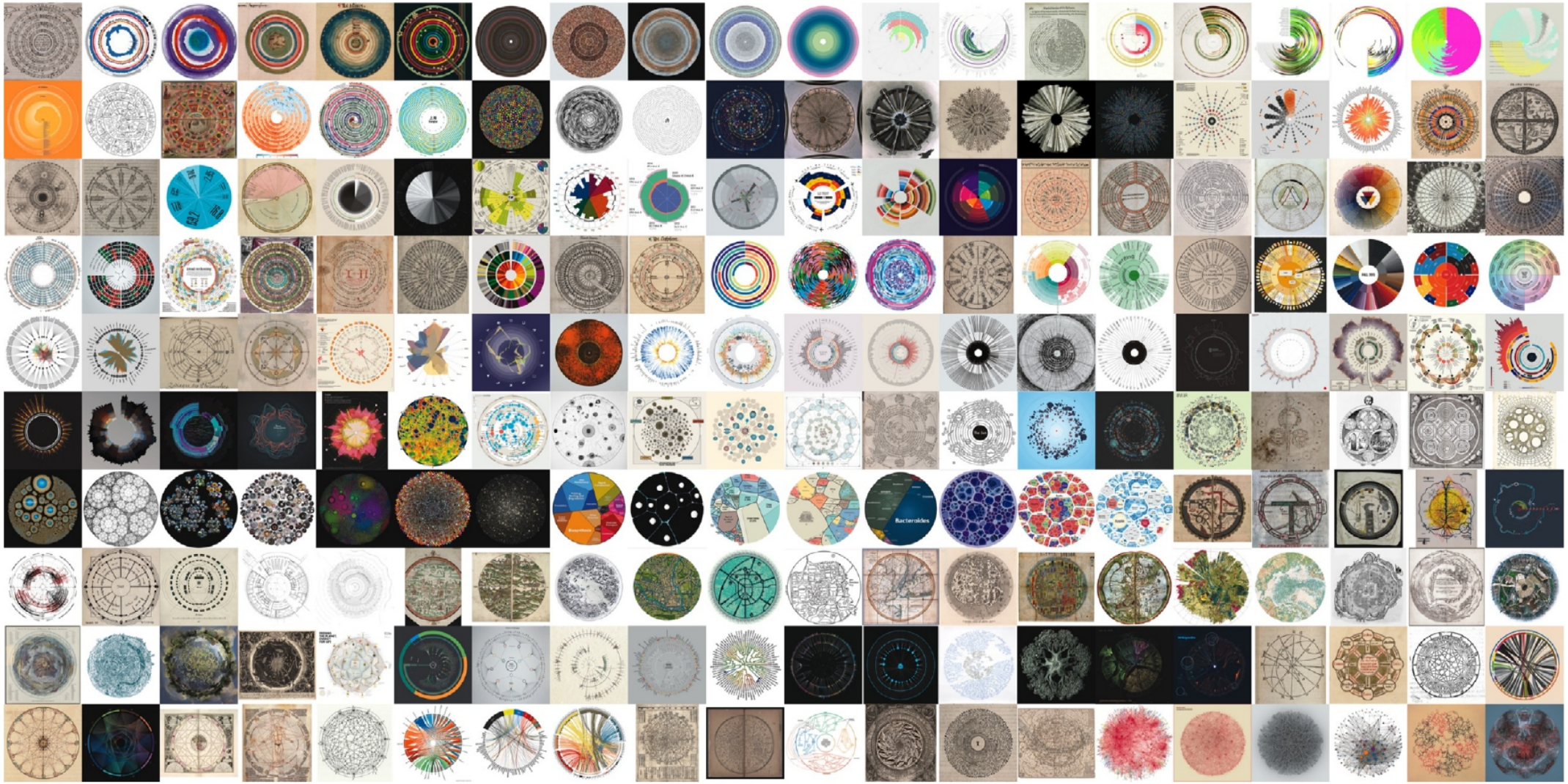
1. Visual Aesthetic
2. Use of color
3. Illustration

Visual Aesthetic
I looked at different aesthetic styles used by visualization designers, they are shown in the next few pages.

Color Palette

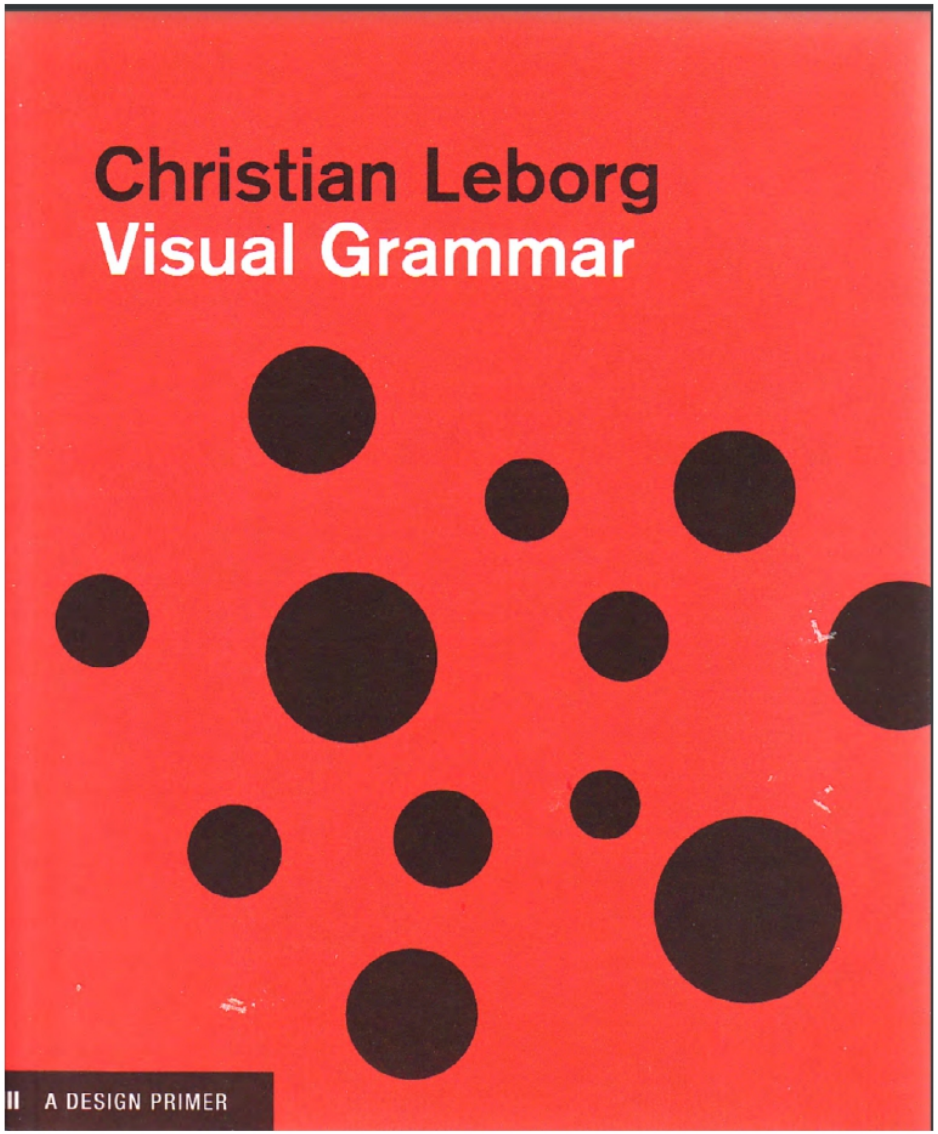
The three main ‘characters’ in the narrative were Finches, seeds, and rainfall. These colors (different tints and shades) were maintained across the different charts and illustrations.

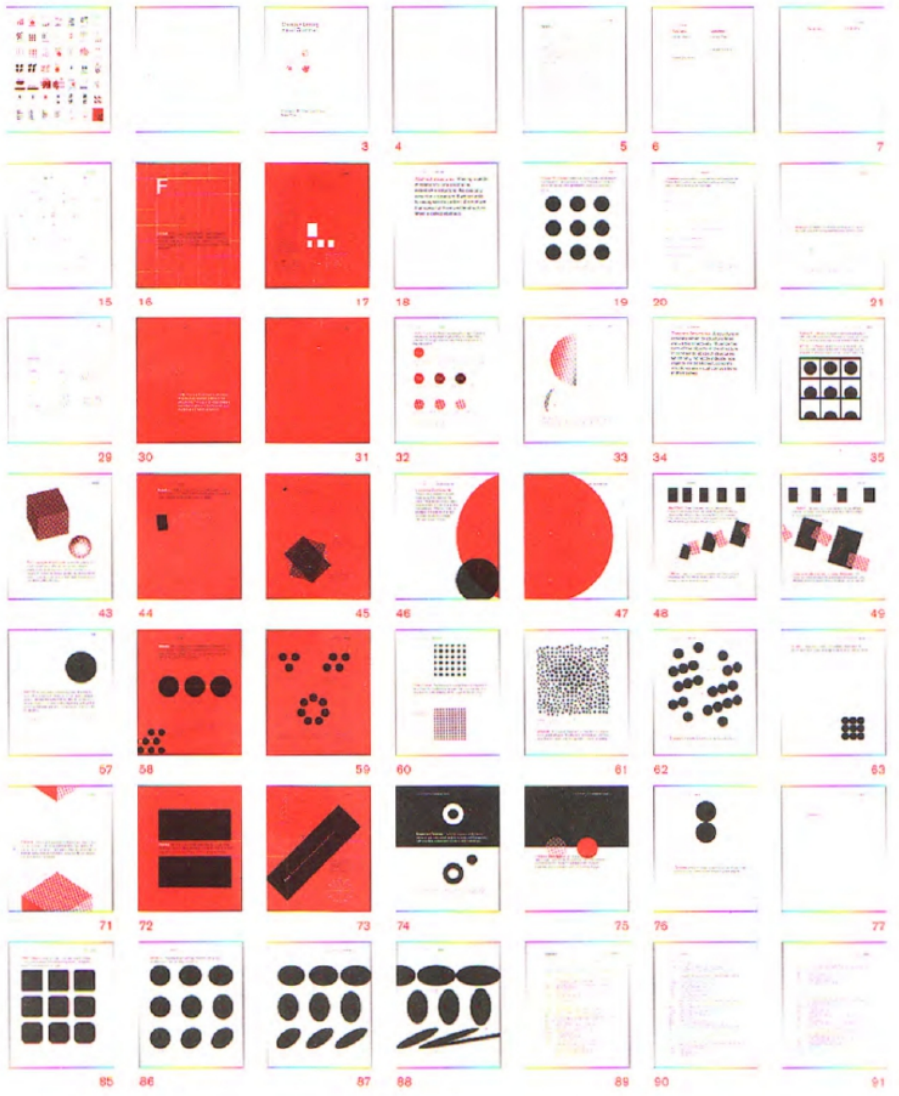




Manuel Lima, 'The Book of Circles:
Visualizing Spheres of Knowledge'
2017

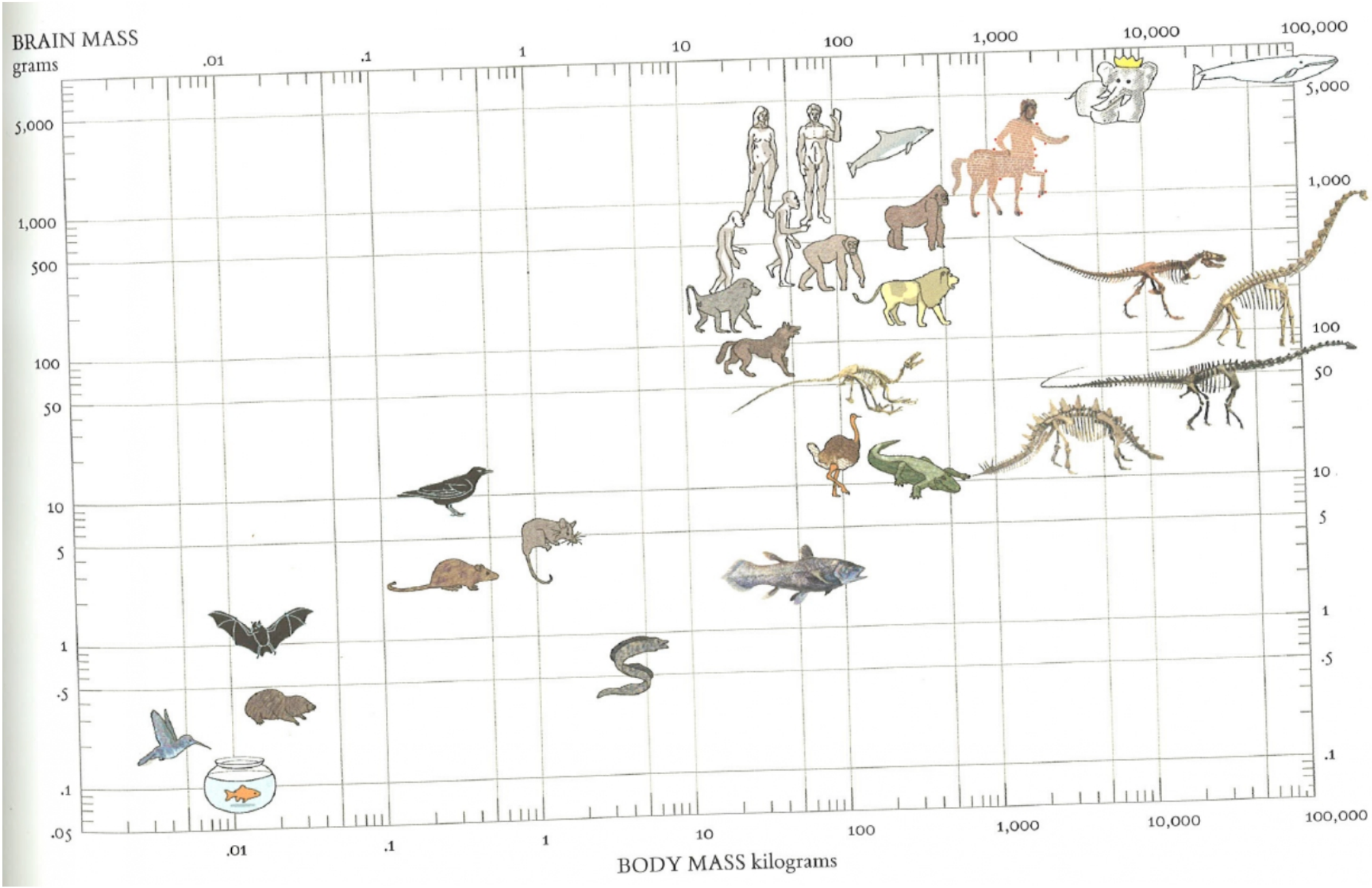
Christian Leborg 'Visual Grammar'
2006

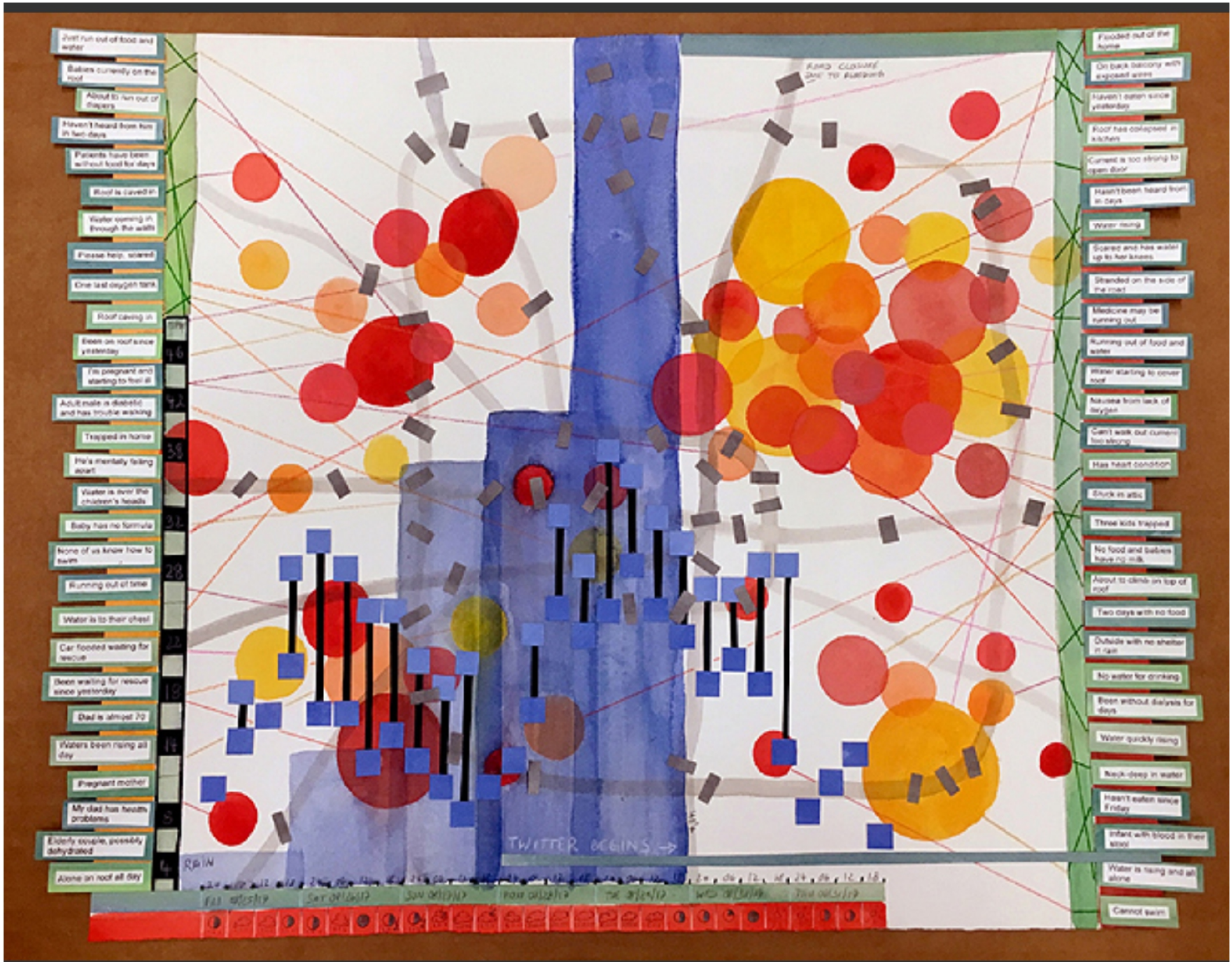




Edward Tufte, Beautiful Evidence (2006) p 121

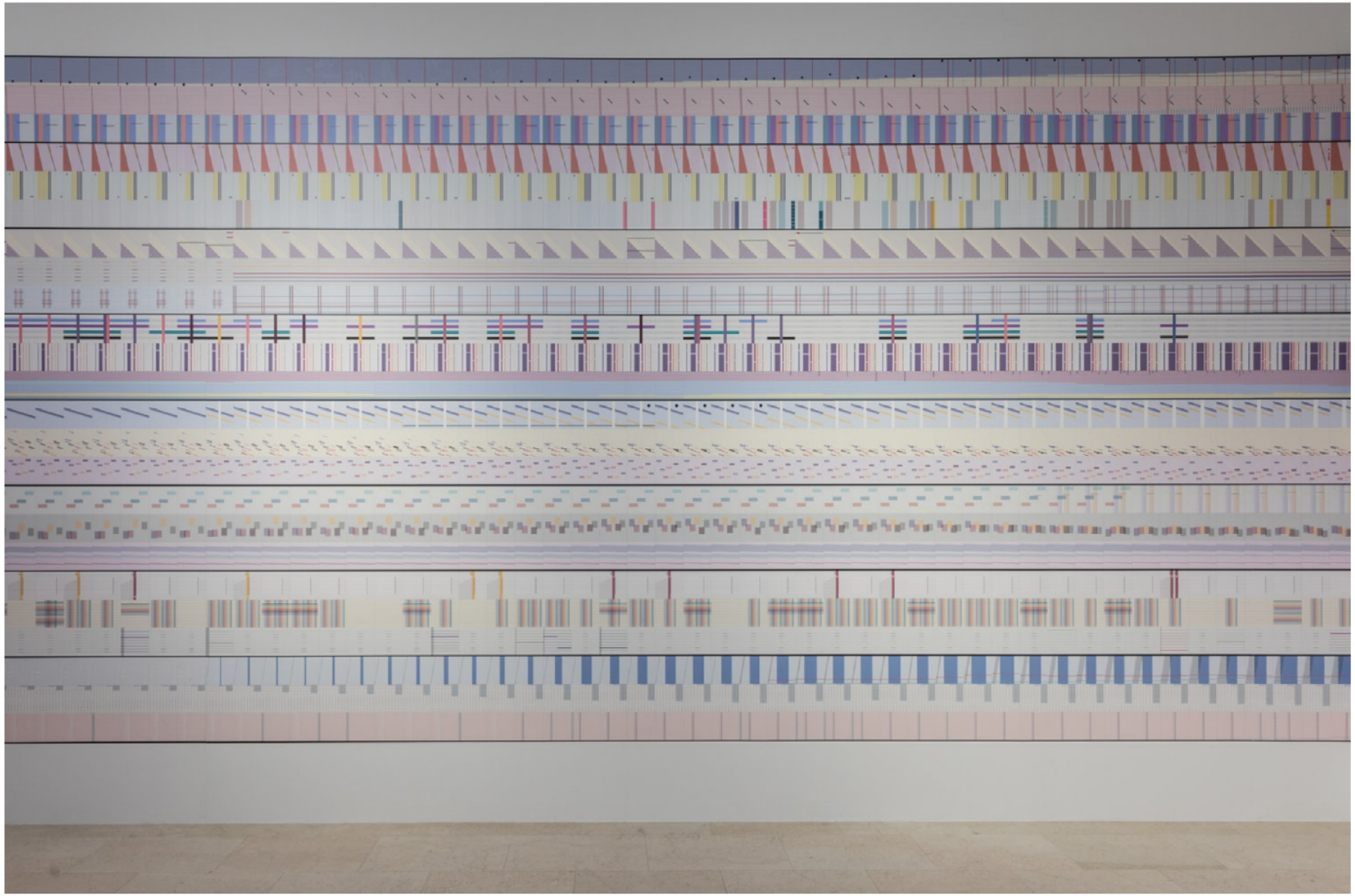
Tufte, E. R. (2006).



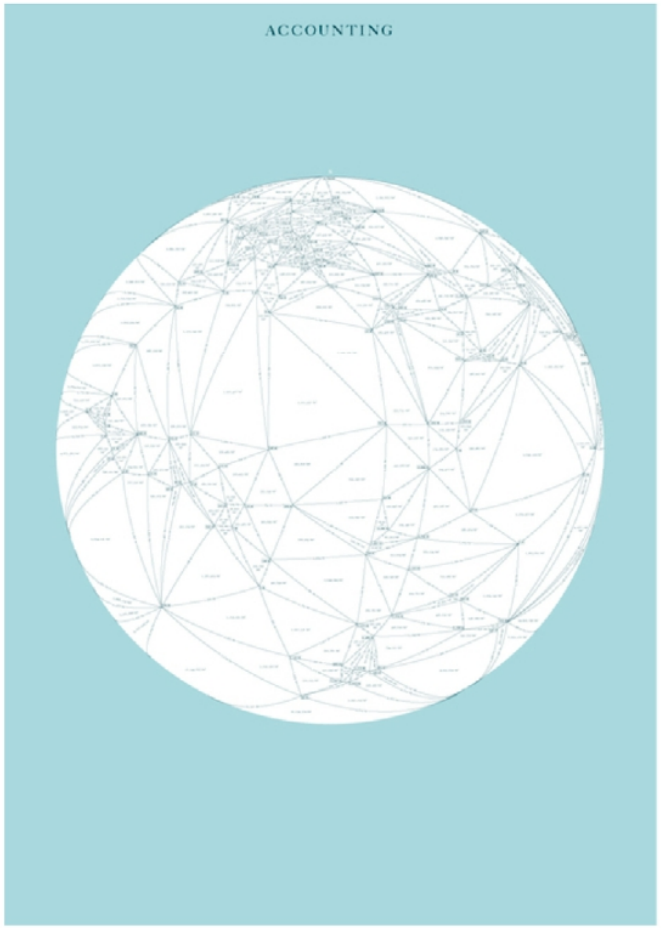


Natalie Miebach 'The Weather score project', 2019

Jer Thorp 'Into the Okavango', 2016



Georgia Lupi, the Room of Change



Stephanie Posavec 'Memory Palace'

Illustration & Icons

McCloud talks about 3 extremes on the picture plane: Reality, visual language, and abstract.

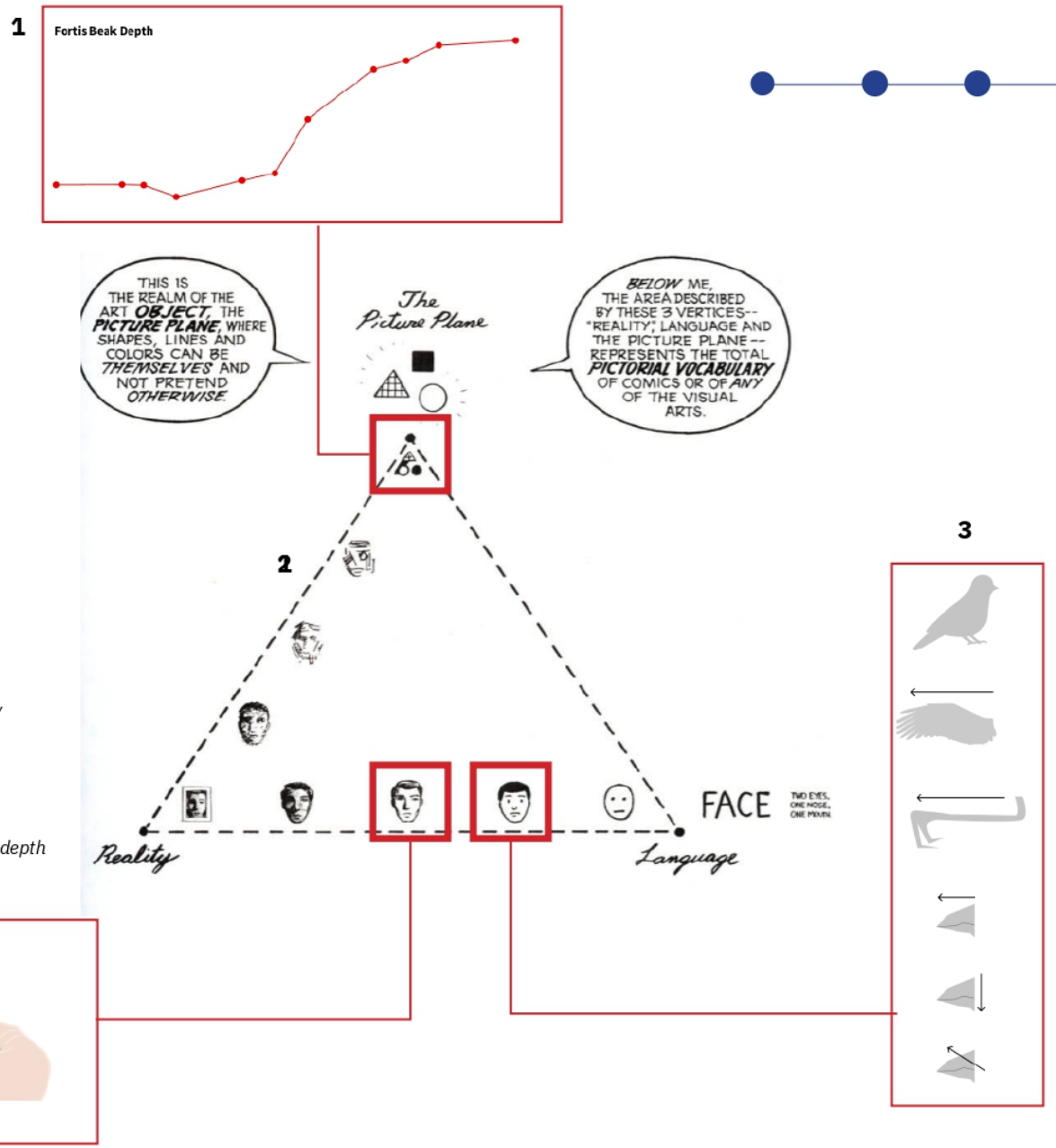
Throughout the data story, I have used 3 types of illustrations/icons, here they are placed on the 'Picture plane'.

(Right) McCloud, S. 1994.

(Below) Illustration for my project by Sukanya Nirmal

(Bottom right) icon used to show measurements of birds

(top) Graph used to show bird beak-depth



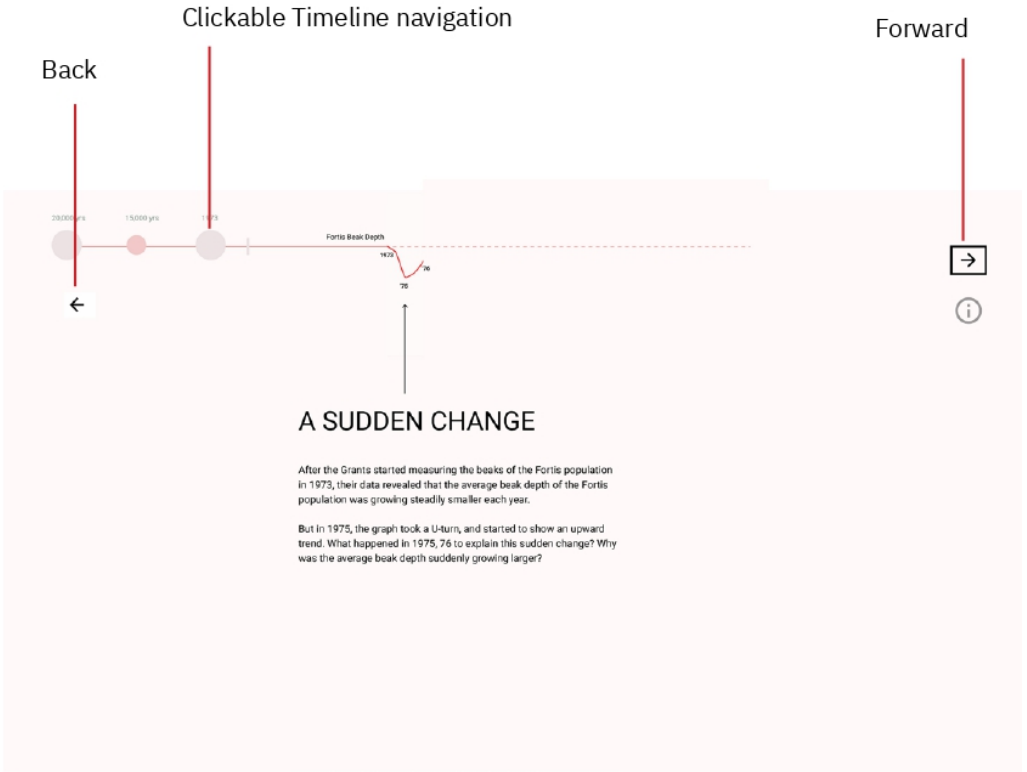
Surface

In this case, the surface refers to the finished product/prototype. The prototyping process would include Interaction and UI Design.
The entire prototyping process was done in Figma. Visualizations were created first in Flourish and then imported into Figma as SVGs.

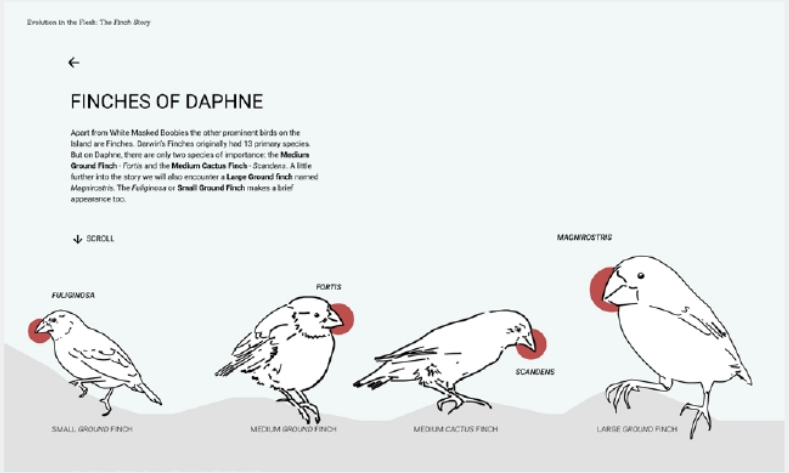
Screen Resolution
1440 x 800
The screen resolution was initially 1440 x 1024 which is the regular desktop resolution. But when tested on smaller laptop screens, the screen height kept getting cut off. After some trial and error, 1440 x 800 worked fine across all users' devices.

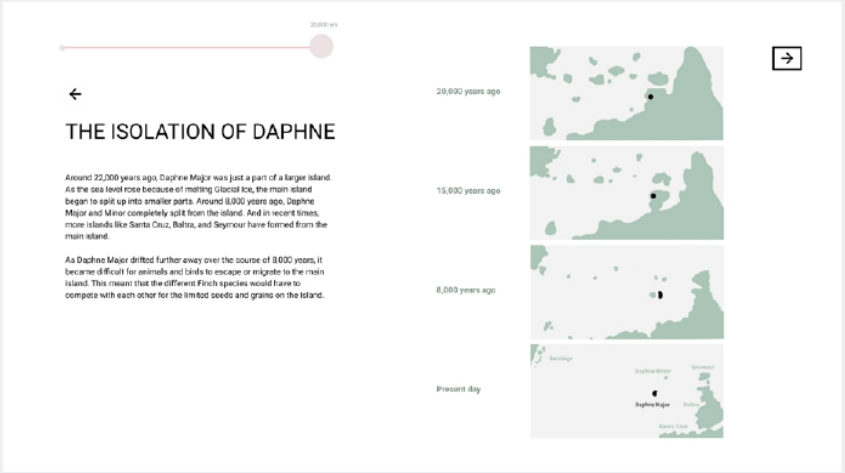
Interaction
The planned interaction was entirely scroll-based, but Figma does not recognize 'scroll' as an interaction trigger. So scrolling was replaced by Navigation buttons.

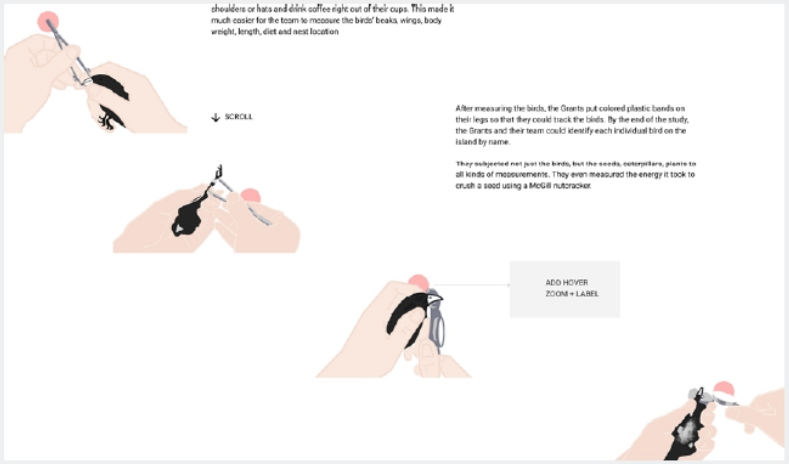
UI Design
While UI design I kept in mind that the interface was not meant to be too complex. It was simply a means to navigate through the story and interact with the visualizations. Navigation initially consisted of back and forward arrows, and the timeline was clickable to move back and forth between events easily.

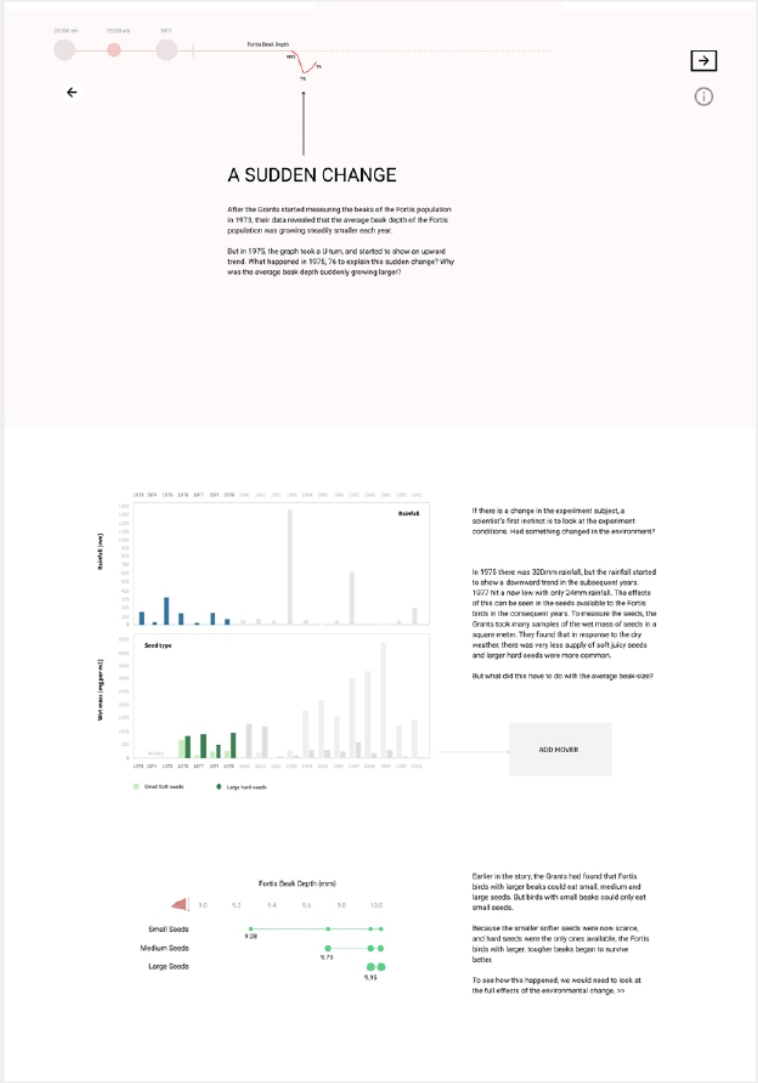


Prototyping









Feedback

I initally showed the prototype to two viewers with different knowledge-levels, interests and reading habits.
One viewer who was not used to focused reading, or interested in Evolution found the text boring, jargon-heavy and hard to follow. He was interested in the hover interactions and asked if the bar-charts could be replaced by something more interesting.
The other viewer who is a focused reader read the entire story and said that it was written quite professionally but the text could be shorter. He also said that the story could be more dramatic to catch the attention of the viewer.

Other than these, I sent the prototype to around 6 people and observed them going through it. Here are some **observations**:

1. In general people paused at the hand-drawn illustrations for longer.
2. Most people did not notice the timeline/pay much attention to it because it was so light.
3. People did not scroll down the pages intuitively and many missed a few pages. This meant that it was not visually communicated on the screen.
4. People commented that the font sizes on the timeline and labels could be made bigger.

Comments:

1. Very nice to look at but the interaction could be more polished.
2. The heirarchy on the pages could be clearer.

Changes

Based on the Feedback I am working to change some of the layouts, cut down text and make the navigation easier by the 26th of December. I will include the prototype with the presentation.

Evaluation

27th-30th Dec
I will conduct a pre and post knowledge-test to evaluate whether the content was communicated clearly. Apart from that viewers will also fill a short survey to gauge interest-level and overall experience.

Future Scope

1. While the scope of this project did not cover the entire study, there are many more interesting insights and discoveries that the Grants made that can be explored further.

2. The project was also done enitrely on Figma which limited the novely to interactions and visualization idioms, however, it can be much more exploratory if it is executed using code.

3. As the previous feedback stated, there is a lot of scope for experimental, unconventional visualisations using this data. That could be in the future scope of this project.

Reflections

One of the greatest challenges during this project was to let go of many interesting facts and details, and ruthlessly scope down the content for the sake of simplicity and flow. Any scientific explanation comes with a fineprint or conditions that might help the user understand just how complex the explanation is. But having too many of those in such a short time would make the text seem erratic and confusing. Also working on this landmark study, and furthermore referring to a book as good as The Beak of the Finch was a great experience. I cannot say that I have added to or improved upon the book in any way. This project is positioned as a gateway to the book and the study.

References

Berinato, S., & Harvard Business Review. (2016, September 14). Visualizations That Really Work. Retrieved October 19, 2020, from https://hbr.org/2016/06/visualizations-that-really-work

Bigelow, A., Druckery, S., Fishery, D., & Meyer, M. (2014). Reflections on how designers design with data. AVI '14 Proceedings of the 2014 International Working Conference on Advanced Visual Interfaces, 17-24.

Brexler, S. J. (n.d.). Public Understanding of Science. Retrieved October 19, 2020, from https://www.apa.org/monitor/2008/12/sd

Cairo, A. (2019). HOW CHARTS LIE: Getting smarter about visual information. Place of publication not identified: W W NORTON.

Echelman, J. (n.d.). Janet Echelman. Retrieved October 19, 2020, from https://www.echelman.com/

Engell, J. (n.d.). Rhetoric: The Art of Persuasive Writing and Public Speaking. Retrieved October 19, 2020, from https://www.edx.org/course/rhetoric-art-of-persuasive-writing-public-speaking

Fathom. (n.d.). Preservation of Favored Traces. Retrieved October 19, 2020, from https://fathom.info/traces/

Gawande, A. (n.d.). What Should Medicine Do When It Can't Save You? Retrieved October 18, 2020, from https://www.newyorker.com/magazine/2010/08/02/letting-go-2

Gladwell, M. (n.d.). How David Beats Goliath. Retrieved October 18, 2020, from https://www.newyorker.com/magazine/2009/05/11/how-david-beats-goliath

Grant, P., & Grant, B. (2014). 40 Years of Evolution: Darwin's Finches on Daphne Major Island. Princeton; Oxford: Princeton University Press. doi:10.2307/j.ctt5hhncc

Hernandez, M. (2020, August 04). In a flash, a changed world. Retrieved October 19, 2020, from https://graphics.reuters.com/WW2-ANNIVERSARY/HIROSHIMA/rlgpdnqljpo/

Leborg, C. (2006). Visual grammar. New York: Princeton Architectural Press.

Lima, M. (2017, March 26). The Book of Circles. Retrieved December 24, 2020, from https://medium.com/@mslima/the-book-of-circles-4b511a5bcffc

Lupi, G. (n.d.). The Room of Change. Retrieved December 24, 2020, from http://giorgialupi.com/the-room-of-change

May, Robert M., & Andrew P. Dobson. (1986). Population Dynamics and the Rate of Evolution of Pesticide Resistance. In Pesticide Resistance. See Roush and Tabashnik, eds. 1986, 170–93.

McCloud, S. (1994). Understanding comics: [the invisible art]. New York:

HarperPerennial.

Miebach, N. (2019). The Weather-score project. Retrieved December 24, 2020, from https://nathaliemiebach.com/weatherscores.html

Munzner, T. (2015). Visualization Analysis and Design. CRC Press. ISBN: 9781498759717

Nichani, Maish, and Venkat Rajamanickam. 2003. 'Interactive visual explainers: a simple classification.'

Rafaeli, Sheizaf, and Fay Sudweeks. 1997. 'Networked interactivity.' Journal of Computer-Mediated Communication 2 (4). Article first published online: 23 June 2006. <http://www.onlinelibrary.wiley.com/doi/10.1111/j.1083-6101.1997.tb00201.x/full>.

Rutstein, R. (2018, June 05). Artist at sea (2015). Retrieved October 19, 2020, from http://rebeccarutstein.com/work/projects-6/

Ryan, Maire-Laure. 2006. Avatars of story. Minneapolis, MN: University of Minnesota Press.

Sims, Rod. 1997. 'Interactivity: a forgotten art?' In Computers in Human Behavior 13 (2): 157–180. <http://dx.doi.org/10.1016/S0747-5632(97)00004-6>.

Thorp, J. (n.d.). Into The Okavango. Retrieved December 24, 2020, from https://www.jerthorp.com/intotheokavango

Tufte, E. R. (2006). Beautiful Evidence. Cheshire: Graphics Press.

Tufte, E. R., & Graphics Press,. (1990). Envisioning information.

Weber, W. (2017). Interactive information graphics - A framework for classifying a visual genre. In Information Design Research and Practice, (1st ed., pp. 243-256). Routledge.

Weiner, Jonathan. (1994). The beak of the finch : a story of evolution in our time. New York :Knopf : Distributed by Random House