

# LEVERAGE RESEARCH HISTORY & FAQ

# Shared Draft

Document Status: Publicly Available Early Draft. Last Updated: January 10, 2023.

Welcome to the Leverage Research FAQ.

Read answers to frequently asked questions about our current work below. For questions about our past work, 2011–2019, see our <u>Leverage Pre-2019</u> FAQ.

This is a living document. We will continue to add and update questions with new information periodically. If a question you would like answered is not included or you think any answers are missing important information, you can provide feedback <a href="here">here</a> or contact us at <a href="here">contact@leverageresearch.org</a>.

#### **Contents:**

About Us: Vision, Mission, Programs, and Approach

<u>Overview</u>

Our Vision

**Our Mission** 

Our Focuses

Why focus on science?

Why early stage science?

Why the social sciences?

Why focus on responsible science?

Why focus on progress?

Our Approach

**Our History** 

**Current Programs** 

History of Science Program

Where can I read your history of science research?

What other case studies do you have planned?

Why are you focusing on the history of electricity first?

Exploratory Psychology Program

What is Belief Reporting?

What is Charting?

What is Connection Theory?

Bottlenecks in Science and Technology

Which Bottlenecks events have happened so far? What events are forthcoming?

What is next for the Bottlenecks Initiative?



### **More Information**

**Donor policy** 

What are your plans for publication?

Where can I find out more?

#### Leverage Pre-2019

The Leverage Project

Basic facts

Major concepts

Major features

Major phases

More Information

**Common questions** 

Why was Leverage 1.0 secret?

Why were people from Leverage 1.0 so optimistic?

Why did Leverage 1.0 dissolve?

What are the major takeaways from Leverage 1.0?

What were the major mistakes made during Leverage 1.0?

<u>Institutional Design</u>

Institutional Operation

Research Program Execution

<u>Internal Culture</u>

Culture, Values, and Policies

What was it like to work at Leverage Research pre-2019?

# About Us: Vision, Mission, Programs, and Approach

### Overview

Leverage Research is an independent, non-profit research institute that seeks high-leverage ways to support responsible scientific advance.

We study and support responsible scientific progress by developing systematic bodies of knowledge in important but underexplored areas which can be used and built upon by others, focusing especially on the early stages of discovery and the role of people in science.

Our three main programs are:

- History of Science: studying how scientific advances occur in nascent fields by conducting
  in-depth case studies into how key discoveries were made in the early history of successful
  fields.
- **Exploratory Psychology:** distributing tools from our early (2011–2019) psychology research to support the external testing of our original hypothesis and further independent research



 Bottlenecks in Science and Technology: coordinating researchers, funders, institution builders, and others to identify technical, social, and institutional bottlenecks to progress in their fields and take effective action to break them.

The institute previously conducted research into a wider range of important but poorly understood topics in the social sciences, such as coordination, practical epistemology, and expert assessment. Much of this research is still unpublished but often informs our efforts and approach today so we dedicate a small portion of our time to releasing research reports on this.

You can find links to our research reports in our <u>Publications and Background Information</u> document. You can read more about our past work, 2011–2019, in the section of this FAQ on <u>Leverage Pre-2019</u>.

### **Our Vision**

We believe a radically better future is possible. We want life to be significantly better for everyone, and believe that it can be with deliberate and careful effort.

We imagine a world where humanity is able to reliably and systematically develop the knowledge and practical know-how to identify and solve the most pressing problems the world faces today—creating abundant clean energy, eradicating diseases, tackling the mental health crisis, and preventing global disasters—and the potential new problems of tomorrow.

The scientific enterprise is perhaps the most successful way of developing new knowledge and new technologies in our species' history. Yet that enterprise has also brought with it new challenges and dangers of its own.

Our vision is of a healthy, responsible scientific enterprise that is equipped to generate and employ the knowledge necessary to identify and tackle the world's most important problems.

### **Our Mission**

Leverage Research's mission is to study and support responsible scientific advance by generating systematic bodies of knowledge in important but under-studied areas which can be used and built upon by others, focusing especially on the early stages of discovery and the role of people in science.

Our approach is to seek out high-leverage opportunities: important areas of research that are under-studied or under-supported where we expect to be able to usefully intervene. We run long-term, foundational research programs aimed at helping build new or struggling fields from the ground up, and take an engineering approach, learning from practical endeavors as much as from theoretical reflection.



The institute focuses especially on the role of people in science, which translates to a special interest in the social sciences, as well as the earliest stages of scientific advance, where the process of discovery is less well-understood.

In our **Early Stage Science program** we are working to create an accessible history of the major discoveries in the early stages of various fields to provide a dataset through which to understand early stage science. In our **Exploratory Psychology program**, we are distributing introspective methods and tools to be used by external researchers to explore the mind, and through our **Bottlenecks in Science and Technology program** we seek to understand the social and institutional bottlenecks to scientific progress and apply our research findings to support efforts in struggling and nascent fields directly.

### **Our Focuses**

While Leverage's work touches on many aspects of science, we focus especially on **early stage science** and **social science**. These are the areas we believe we can have the largest impact, both because these areas are neglected and important, and because our unique approach and skillset position us especially well to be able to understand both.

With respect to all aspects of our work, we focus on the twin goals of **progress** and **responsibility**. In our view, these goals are intimately related. Irresponsibility within science will both harm people science should be supporting and ultimately also yield slower progress. Yet, at the same time, making progress is part of humanity's responsibility to solve the challenges it faces.

# Why focus on science?

Our vision is of a world where groups that need knowledge have the means and capacity to develop that knowledge when they need it. In the world of today, many pressing questions remain unanswered and are not on track to be answered in a timely fashion.

With so much still to be known, and a possibility space so vast, it makes sense to specialize. Leverage Research has chosen to focus on aiding the development of new *scientific* knowledge, rather than other types of knowledge (such as mathematical or philosophical), because of the astonishing power and prosperity science and technology have brought over the last several centuries. If humankind is to acquire the abundance of resources it needs and equip itself to enjoy that abundance, it will in part be because humanity has learned to develop new scientific knowledge more reliably and more responsibly.

Recommended reading:

The impact of Science on Society by James Burke, Jules Bergman, and Isaac Asimov.



### Why early stage science?

We focus primarily on early stage science for three reasons.

Firstly, we believe it is possible that the early stages of discovery may be critical in the development of fields and that progress in the early stages looks very different from science in the later stages, but popular conceptions of science are based primarily around these later stages. We write about this hypothesis in our <u>program introduction</u> on early stage science.

Secondly, we believe that it is a high-leverage research avenue. How to make progress in the early stages of a field is understudied and commonly misunderstood, but important, especially for nascent or struggling fields. Our current hypothesis is that one reason many social sciences have not enjoyed the success of the hard sciences might be that misunderstandings of the history of successful scientific fields lead researchers and funders to aim to replicate successful late stage science, rather than helping the social sciences through the early stages that the hard sciences went through.

Finally, early stage research is an area we have some experience with having previously studied epistemology, experimented with different research program designs, and conducted our own experimental, early stage research attempts. We therefore believe that it is possible to make progress and believe we are well placed to do so.

Read more about our history of science research in the <u>History of Science Program</u> section of this FAQ or on the <u>History of Science program page</u> on our website.

# Why the social sciences?

While fields like chemistry and physics have flourished, many in the social sciences worry that their disciplines have failed to progress as hoped. Researchers in the "soft" sciences fear that they are not measuring up to their "hard" counterparts, and discussions of replication crises have become commonplace in psychology, economics, and a host of other areas. This is not because of importance — understanding people (the object of the social sciences) is essential for solving many of the world's problems, and also may yield insight into the research process itself.

<sup>1</sup> See, e.g., Daniele Fanelli, "Positive' Results Increase Down the Hierarchy of the Sciences," *PLOS ONE* 5, no. 4 (April 7, 2010): e10068, <a href="https://doi.org/10.1371/journal.pone.0010068">https://doi.org/10.1371/journal.pone.0010068</a>; Jennifer L. Tackett et al., "Psychology's Replication Crisis and Clinical Psychological Science," *Annual Review of Clinical Psychology* 15, no. 1 (2019): 579–604, <a href="https://doi.org/10.1146/annurev-clinpsy-050718-095710">https://doi.org/10.1146/annurev-clinpsy-050718-095710</a>; Abel Brodeur, Nikolai Cook, and Anthony Heyes, "Methods Matter: P-Hacking and Publication Bias in Causal Analysis in Economics," *American Economic Review* 110, no. 11 (November 2020): 3634–60, <a href="https://doi.org/10.1257/aer.20190687">https://doi.org/10.1257/aer.20190687</a>.



We believe supporting social science research is a high-leverage area for us in particular because we have previous experience conducting research in related areas (sociology, psychology, and coordination), and because we believe there are implications from our history of science research for the social sciences. We hypothesize that many social sciences have not enjoyed the success of the hard sciences because, in order to be considered scientific and receive funding, they are mistakenly encouraged to emulate methods of successful late stage science. It is possible this is a mistake based on common misconceptions about how discoveries were made historically in the early stages of different sciences.

Read more about our previous social science research, as described in the history of our Exploratory Psychology Program in this FAQ and on the Exploratory Psychology program page on our website.

### Why focus on responsible science?

We emphasize responsibility in the context of scientific advance to remind ourselves that while science has provided untold benefits, both materially and in how we understand the world, how these developments are made and their impact on society and individuals matters. Scientific advances are not without significant risks and dangers; splitting the atom provided nuclear energy but also nuclear weapons that have placed the world itself in danger while other research avenues like gain of function and artificial intelligence research have risks that may outweigh their promise.

We believe that making sure scientific and technological advancement is approached responsibly will not only make humanity safer, but also contribute positively to the pace of progress.

# Why focus on progress?

The benefits from scientific and technological progress over the last several centuries are obvious. What may be less obvious are the sources of that progress.

We believe that in many cases, progress in science and technology is not automatic, but instead arises from planning and deliberate effort. This means that to ensure that there is continuing progress in many key areas, purposeful action may need to be taken to cause or support it.

# **Our Approach**

We believe that, as an independent research institute, it is important to think carefully about how we can make the most of the resources we have and the support we receive. Our approach is to look for promising ways to do things differently; seeking to leverage (hence the name) neglected areas of knowledge or related opportunities to understand and improve the world.



We tend to prioritize areas where important problems and promising avenues exist but are currently underexplored. We believe that there are a lot of valuable research avenues available that are often overlooked for social or structural reasons. We see people, and the proper application of human thought, ingenuity, and judgment, as central, leading us to a greater focus on the social sciences as part of our research.

As an independent research institute, some of the ways in which we are able to do things differently include:

- 1. **We conduct foundational research:** Research progress often requires sustained, long-term efforts to identify and tackle the key issues which are further upstream of specific problems. When we select a research area, we aim to develop a picture of how that field works on a technical, social, and institutional level. This frequently leads us to identify opportunities for approaching problems in the field differently. Depending on how novel the approach, this sometimes results in our trying to rebuild a field from the ground up. Such efforts require deliberate design and may take many years.
- 2. **We take an engineering-centric approach:** In many cases, the earliest stages of research take place more effectively if they are conducted in concert with an attempt to build or do something in the relevant domain, rather than by theoretical reflection or hypothesis-driven experimentation alone. As a result, we frequently select concrete goals in domains we are studying and attempt to 'learn by doing' as we seek to achieve those goals, developing theories and engaging in hypothesis testing along the way.
- 3. We have freedom to explore novel research avenues: Valuable ideas can often come from unusual or overlooked knowledge traditions or areas of research that have not yet reached the mainstream. Yet, following such avenues can be difficult due to fear of pushback from a variety of sources. The culture of exploration we develop and donors who believe in the value of our work affords us the freedom to explore less traditional avenues of research where necessary to make progress.

# **Our History**

Leverage Research was founded in 2011 as an experiment in research and coordination. From 2011 through 2019, we sought to create an environment where talented, agentic individuals could work on projects intended to meaningfully impact the world, either designing their own or working on projects designed by others, all as part of a largely decentralized research collaboration.

During the eight and a half years of this experiment, now sometimes referred to as "Leverage 1.0," the research project grew to encompass 45 collaborators, spanning multiple organizations. The primary focus ended up being research on topics in the social sciences, though important effort went into other projects as well. By 2019, coordination among sub-teams had broken down, and in



mid-2019, the research collaboration was dissolved with most team members moving on to other projects.

Leverage Research then re-organized and professionalized, adopting a new set of programs and narrowing its mission. This new phase of the institute's life is sometimes called "Leverage 2.0," to distinguish it from the more freewheeling days of its past.

Controversy that sometimes exists around Leverage Research is almost exclusively centered around Leverage 1.0 and the variety of unusual approaches and strategies we adopted, both as we bootstrapped ourselves into existence and as we sought to research a surprisingly difficult set of questions on topics pertaining to human nature.

To learn more about Leverage 1.0, see the section of this FAQ on Leverage Pre-2019.

# **Current Programs**

Leverage Research current has three main programs:

- History of Science: studying how scientific advances occur in nascent fields by conducting in-depth case studies into the factors behind key discoveries in early history of successful fields.
- **Exploratory Psychology:** distributing tools from our early (2011–2019) psychology research to support the external testing of our original hypotheses and further independent research.
- Bottlenecks in Science and Technology: coordinating researchers, funders, institution builders, and research program managers to identify technical, social, and institutional bottlenecks to progress in their fields and take effective action to break them.

Our programs involve new foundational research, the dissemination of our previous research, or an attempt to enable others to benefit from the application of our research. In our Bottlenecks program, for instance, we are bringing to bear knowledge acquired from our previous study of research program design and human coordination as well as our current study of the history of science.

# **History of Science Program**

How were discoveries made in the past, what did scientific progress look like in the early stages of today's most successful fields, and what are the most important factors that contribute to progress?



How one thinks of science and its history exercises a powerful influence over one's approach to many areas of public concern. Which narratives get told and what's taken as progress guide decisions about which projects to fund and which to terminate; they influence how results are communicated and received, which standards are applied to which field, and what discoveries are taken up as breakthroughs. Directly or indirectly, they shape nearly every aspect of policy in which empirical findings play a part.

Yet the vision of science brought to bear on these pressing issues is, as a rule, quite partial. As much as the tendency has been opposed by historians and philosophers of science, the dominant understanding of research is as a theory-driven (and frequently quantitative) enterprise. The models against which developing areas of inquiry are assessed come primarily from established, late-stage sciences such as physics and chemistry, and at the level of individuals, the prevailing image of the scientist remains closer to Einstein than to Faraday or Jane Goodall (though the latter are every bit as legitimate as exemplars of the type). As productive as comparisons to the exact sciences can be for younger or less-established fields, however, they are misleading when made in isolation.

The purpose of our history of science program is to develop a more complete picture. By focusing on the early stages of presently successful sciences, we believe that important lessons may be drawn for new or struggling areas of inquiry. We are thus working to produce a catalogue of detailed case studies describing early advances in established fields and the social, technological, and personal factors that allowed each to be made. In the long-term, our aim is to provide enough accessible, historical data to allow a balanced analysis of the circumstances and methods conducive to discovery.

# Where can I read your history of science research?

You can find all of our case studies on our research page: https://www.leverageresearch.org/research.

Each of our case studies has a dedicated page where you can find the full case study and a summary of research highlights detailing the key findings from the case. Since we began this research program in late 2019, we have written the following case studies:

- William Gilbert and the Discovery of 'Electricks'
  - Webpage
  - Case Study
  - Research Highlights
- The Reception of Volta's Electrophorus Among Eighteenth-Century Electricians
  - Webpage

<sup>&</sup>lt;sup>2</sup> See, e.g., Ian Hacking, *Representing and Intervening*, vol. 279 (Cambridge University Press Cambridge, 1983); Friedrich Steinle, "Entering New Fields: Exploratory Uses of Experimentation," *Philosophy of Science* 64 (1997): S65–74.

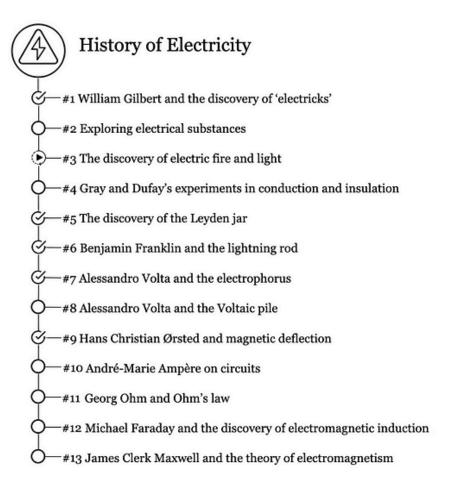


- Case Study
- Research Highlights
- The Discovery and Impact of the Leyden Jar
  - Webpage
  - Case Study
  - Research Highlights
  - <u>Kleist Letters:</u> three previously unpublished letters from Kleist detailing his early studies, which we transcribed and translated as part of our research.
- o Ørsted and the Discovery of Electromagnetism
  - Webpage
  - Case Study
  - Research Highlights
- The Development and Spread of Franklinian Theory
  - <u>Website</u>
  - Case Study
  - Research Highlights

# What other case studies do you have planned?

Ultimately, we would like to produce case studies on all of the major discoveries in the early history of the successful sciences, including astronomy, magnetism, chemistry, optics, geology, and the life sciences. For now, our efforts are concentrated on the history of electricity which, according to our most recent estimate (given in our 2021 Annual Report), will require thirteen case studies. The following diagram displays what those cases are and our progress with respect to each:





# Why are you focusing on the history of electricity first?

We initially chose to focus on a single field, rather than doing case studies across multiple fields, because we thought this would make it easier to compare case studies as we worked out our case study format. We chose the history of electricity because it seemed interesting, complex, and understudied.

Having now done a number of case studies in the history of electricity, we expect to prioritize finishing the rest as this will make it possible to make judgments about the early history of an entire field. That said, we have some degree of flexibility and may decide to begin case studies in other fields if we encounter good reasons to do so.

# **Exploratory Psychology Program**

What can introspection teach us about the human mind? Is it possible to create a new flourishing subfield of psychological research based on new introspective tools and methods?



Modern-day psychological research takes many forms. Despite the variety, most such research tends to be data-driven rather than theoretical, centered around hypothesis testing rather than engineering, and focused on third-personal rather than introspective methods.

From 2012–2019, Leverage Research conducted a theoretically-informed, engineering-centric, introspective investigation of the mind. This research led us to develop a large number of hypotheses about the content, structure, and functioning of the mind. Our Exploratory Psychology Program aims to disseminate our research methods, equipping others to conduct their own investigations and test our hypotheses.

To learn more, visit our <u>Exploratory Psychology Program page</u>.

### What is Belief Reporting?

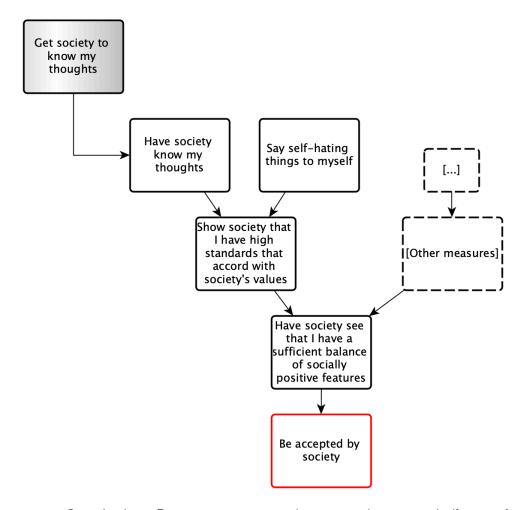
Belief reporting is an introspective method that allows a person to test whether they believe a proposition or its negation. We developed it in 2014, after having examined existing introspective methods and found them inadequate for our research needs. Belief reporting figured heavily in our psychology research from 2014–2019, and is one of the research tools we will be distributing as part of our Exploratory Psychology program.

Belief reporting has some similarities to <u>Focusing</u>, an introspective method developed by <u>Eugene</u> <u>Gendlin</u> and presented in his 1978 book *Focusing*. For parties interested in an introduction to introspection, we recommend the *Focusing* <u>audiobook</u>.

# What is Charting?

Charting is a method for organizing introspective reports. When written down, charts take the form of diagrams composed of boxes and arrows, where the boxes include belief content and frequently represent goal structure.





Sample chart. Represents one mental structure that can underlie negative self-talk.

The original version of charting was developed by Geoff Anders in 2008 as part of testing <u>connection theory</u>. It was then improved and refined by several researchers as part of Leverage's psychology research during 2012–2019. Charting is now often used in conjunction with <u>belief reporting</u>, and is one of the research methods we will be distributing as part of our Exploratory Psychology program.

# What is Connection Theory?

Connection theory (CT) is one of the research tools employed by our researchers during our early psychology research in 2012–2017, and then in our intention research in 2018–2019. As part of our Exploratory Psychology program, we expect to distribute information about how to use connection theory as part of introspective research.

CT is a theory of human belief and action, developed by our founder in 2008. In combination with a few auxiliary claims, CT is a framework that can be used to guide introspective investigations, make predictions about what will cause changes in a person's beliefs and actions, and design interventions meant to help people make changes in their beliefs or actions.



# **Bottlenecks in Science and Technology**

How much scientific and technological progress is being made, what are the most important bottlenecks to further progress, and what is required to break those bottlenecks?

Over the past decade, we've heard narratives of scientific and technological progress both accelerating (singularity) and slowing down (stagnation). Our Bottlenecks in Science and Technology program asks the question: How much progress is really being made in each field? Why isn't progress being made faster? What are the social, institutional, and technical bottlenecks to further progress, and how can those bottlenecks be broken?

As part of this program, we are launching the Bottlenecks Initiative, a coalition of researchers, funders, institution builders, and thinkers interested in supporting responsible scientific progress by identifying the bottlenecks to progress in different fields and developing actionable plans to break those bottlenecks.

Leverage Research's role in this network is to help the individuals and projects involved by running events, connecting researchers and funders, supplying expertise, and applying our institutional knowledge to help teams overcome the bottlenecks they identify.

# Which Bottlenecks events have happened so far? What events are forthcoming?

Our first Bottlenecks event was a workshop entitled "Bottlenecks in Science and Technology," aka "Bottlenecks 2021," which we co-organized <u>José Luis Ricón</u> and <u>Adam Marblestone</u> in June 2021. The workshop was hosted at the headquarters of <u>Boom Supersonic</u> in Colorado, with keynotes from Peter Thiel, Patrick Collison, and Tyler Cowen and presentations from researchers on the bottlenecks in ten different fields. To learn more about the workshop, see this <u>summary</u> from one of our co-organizers, or read our Q2 2021 newsletter on our <u>updates page</u> and our <u>workshop review document</u>. A few talks from the workshop are available on the <u>Bottlenecks YouTube Channel</u>.

A few months later, some participants from the original Bottlenecks workshop together with <u>Prima Materia</u> hosted a TEDx-style Bottlenecks event entitled Bottlenecks: London.

In 2022, we look forward to running Bottlenecks 2022 and hiring an <u>Events and Operations Manager</u> to expand our capacity to run other events as part of this program.



### What is next for the Bottlenecks Initiative?

In 2022, we plan to organize the construction of roadmaps for scientific and technological progress in a variety of fields, run further events, and continue to build the Bottlenecks network.

To learn more, visit our <u>Bottlenecks in Science and Technology program page</u>.

### More Information

# **Donor policy**

During 2012–2019, we sought to make rapid research progress on many topics in the social sciences, communicating little with the public. During this time we were primarily supported by donations from a small number of tech entrepreneurs who believed in our mission, team, or founder.

After 2019, the institute adopted the goal of communicating its results to the public and building a broad base of support for its work. This led us to adopt, in 2021, a self-imposed limitation on the size of donations: \$10,000 per year from individuals and \$50,000 per year from organizations. Through this limitation we hope to set up our incentives properly and keep the size of the institute in step with the public's understanding of our work.

In rare cases, we accept larger donations from organizations or individuals who share our perspective on the scientific and technological landscape, our approach to driving progress, and support our independent decision-making. At present, the only donor in this category is the Thiel Foundation, with whom we share the values of responsible scientific advance and bold, independent action. The Thiel Foundation is donating \$135,000 per year for the next few years, which amounts to less than 1/3 of our projected annual budget.

We expect that donors to Leverage Research will understand its propensity for agentic action on the basis of its own <u>strategic picture</u>, and as such endorse our having the freedom to pursue our mission as we see fit.

We value donations not only as a source of funding, but also as a visible sign of support and as a source of motivation and inspiration for our team. If you would like to support us financially, please visit <a href="https://www.leverageresearch.org/donate">https://www.leverageresearch.org/donate</a>.

# What are your plans for publication?

We believe that publication is an important step in the research process, with different plans for publication for different parts of our research.

History of science. Our history of science case studies are written for publication on our website. As we complete these studies, we sometimes encounter important missing pieces in the academic



literature. In such cases, we write pieces derived from our case studies for publication in relevant journals. Thus far we have had two essays accepted for publication, one in *Isis* and the other in *Lias*.

In some cases, we also write brief historical pieces that help to give context to our other research. In such cases, we may publish in relevant journals as the occasion arises. For instance, as of August 2022, we have one piece under review at <u>Seeds of Science</u>.

Exploratory psychology. Once we have helped to create an external research community that experiments with our psychology research methods and have collected enough data, we expect to publish reports on the results. These reports will be written primarily for our website, though we also hope to find researchers interested in our work who are well-situated to publish in academic journals.

*Previous research.* As time permits, we are writing up parts of our previous research in the form of research reports. These are currently available on our <u>Publications and Background Information</u> document on Medium.

### Where can I find out more?

You can learn more about the current work of Leverage Research in the following places:

- Leverage Research website: <a href="http://leverageresearch.org/">http://leverageresearch.org/</a>
- Leverage Research Medium: <a href="https://medium.com/@LeverageResearch/">https://medium.com/@LeverageResearch/</a>
  - **Medium series:** Introduction to Leverage Research
- Our quarterly newsletter: <a href="https://www.leverageresearch.org/updates">https://www.leverageresearch.org/updates</a>
- **Twitter:** <a href="https://twitter.com/leverageres">https://twitter.com/leverageres</a>
- Facebook\*: https://www.facebook.com/LeverageResearch/
- **LinkedIn\*:** https://www.linkedin.com/company/leverage-research

For a more complete list of resources see our <u>Publications and Background Information</u> Medium post.

If you have questions that are not covered by this FAQ, you can submit them using the form on our website: <a href="https://www.leverageresearch.org/questions-feedback">https://www.leverageresearch.org/questions-feedback</a>

\*we're not currently that active on Facebook and LinkedIn, though we try to post occasional updates.



# Leverage Pre-2019

During its first eight and a half years, Leverage Research ran a research project that we sometimes refer to as "Leverage 1.0." The project was unique in aim, structure, methods, and outcomes, an object of intense focus for insiders and fascination for outsiders.

The Leverage 1.0 project was so different from the institute today, and so different from most research organizations people have typically encountered, we believe it deserves its own FAQ. In fact, many of the questions we receive about Leverage today are actually questions about Leverage 1.0. Now that the project has completed, and given that—unlike the Leverage 1.0 collaboration—today we are much more focused on public communication, we believe it is possible to answer many of the questions we've received over the years.

This FAQ is a living document, which we will add to and update over time. If there's something you'd like to know or understand about Leverage 1.0, or Leverage Research's work between January 2011 through June 2019, please submit questions <a href="mailto:here">here</a> or contact us at <a href="mailto:contact@leverageresearch.org">contact@leverageresearch.org</a>.

In some cases, information below is pieced together from documents, notes, interviews, and memory. As such, some information may be updated over time.

# The Leverage Project

### **Basic facts**

Here are some of the basic facts about the original Leverage project:

Dates: January 1, 2011–June 30, 2019

Duration: 8.5 yearsPeak size: 45 people\*

• Effort: 140 person years\*

• Cost: \$3.8m\*\*

<sup>\*</sup>Close estimates. More precise figures may be available later.

<sup>\*\*</sup>This covers the non-profit organizations that were part of the research collaboration from January 2011 through June 2019. The for-profit organizations paid for themselves.



# **Major concepts**

The Leverage project consisted of three things which, when understood together, yield a picture of the project as a whole.

Leverage 1.0 was:

- (1) An open research center. Starting in 2011, we set out to create an open, dynamic research environment by hiring the most ambitious, thoughtful people we could and putting them all in the same place. Our belief was that by bringing together the right people and giving them the freedom to choose what to work on, we could produce more knowledge more quickly than by any other means.
- (2) A stealth attempt to pull off the human self-improvement cycle. Early in 2011, and again in early 2012, some of our researchers suggested that if we could make enough progress understanding people, we could use that knowledge to make ourselves better, and in particular better at understanding people. With enough effort, we thought, it might be possible to spark a positive feedback loop where our advances would make us better able to make advances of the same and different types. In April 2012, we concluded that some of this research should be conducted under conditions of partial secrecy. (For more on this, see below.)
- (3) An attempt at heroic responsibility. Leverage's founder and many subsequent team members tried to adopt an attitude of heroic responsibility. This meant trying to take as much responsibility for the problems in the world as possible. Accordingly, no problem was considered definitively too hard or permanently out of scope.

Many of the surprising and unusual features of the Leverage project arose from trying to do these three things at the same time.

# **Major features**

Here are some of the features of the Leverage project. We expected to add more to this list over time, especially as we get feedback on which parts of the project remain difficult to understand.

Decentralized structure. Part of giving researchers freedom to work on what they wanted
also involved giving them freedom to start new projects and structure those projects as they
saw fit. The result was a decentralized system with a shifting and changing sets of teams
and initiatives. For the first few years, it was difficult for everyone to keep track of what was
happening; we solved this problem in 2014 by putting names of teams and pictures of
members of teams on the wall in the main conference room, updating it as the project
changed and the wall came to be out of date.



- Anticipated change. As an organization attempting to take heroic responsibility, and very aware of our own ignorance, we expected that in the success case we would learn a very large amount about the world and change a lot as a result. The degree of change we needed to plan for increased as we got deeper into the task of trying to self-improve. In concrete terms, this meant that we expected that our view of the world, our research methods, and the organization itself would change dramatically over the course of time, in some cases on a year-by-year basis.
- Special hiring criteria. We anticipated a large amount of change both in terms of our knowledge of the world and in terms of ourselves personally, and needed people who would be likely to stick with the project through those changes on a rational basis. We also anticipated the need to take agentic action at various points, which might produce external opposition. This led us to develop special hiring criteria, aimed at selecting individuals who could persist despite opposition, adapt to potentially large changes in worldview, and flourish in an environment of intellectual disagreement. This narrowed the recruitment pool substantially, ruling out many otherwise talented people.
- Active recruiting. With a narrowed recruitment pool, Leverage worked extremely hard to
  find people who were a good fit, actively searching through many communities and sending
  staff on international recruiting missions. Our modus operandi was not to announce open
  positions and wait for people to come to us, but instead to find people who might be
  interested in the project, wherever they were.
- Information management. Once on the project, it was essential that people both be able to engage in an open environment with many projects of different types and be able to manage information discreetly. We also needed to retain our ability to recruit, despite secrecy. This led us to design information management practices that would enable us to share information without it being likely to spread further on its own. This led to an extremely confusing pattern of communication, especially to the communities from which we recruited; the effects of this are still a source of controversy today.

For more on this, see the section on Why was Leverage 1.0 secret?

• Learning from scratch. Many of the things we were doing were new, or needed to be approached from new angles. Simultaneously, because of the special hiring criteria and narrowed recruiting pool, the best candidates we found frequently lacked desirable skills and experiences. As a result, in many cases we had to learn to do things ourselves. This dovetailed with our focus on self-improvement, but also meant that in many cases we were trying to do things we did not know how to do.



# Major phases

Leverage 1.0 went through several important phases. These could be broken up differently depending on one's purpose, but for the overall project it is reasonable to consider it to have gone through four key phases:

**Phase #1: Exploration.** From January 2011 until June 2014, we explored a large number of potential avenues for impact. In accordance with our effort to take heroic responsibility, we considered nothing off limits. In practice, we especially explored avenues that involved developing new knowledge and new technologies, including mental technologies (e.g., meditation techniques) and social technologies (e.g., consensus practices).

During this time, we ruled out a number of research focuses, including artificial general intelligence, brain- or pharmaceutical-based intelligence amplification, and sleep optimization. *Prima facie*, these seemed like promising opportunities for impact, especially the ones that might yield positive feedback loops. However, we found them alternately too risky (e.g., artificial general intelligence), too unlikely to work in the near term (e.g., pharmaceutical-driven intelligence amplification), or too difficult relative to the expected payoff (e.g., sleep optimization).

**Phase #2: Consolidation.** By June 2014, our research had identified psychologically-informed training as a highly promising route, both for helping people and for yielding powerful self-improvement tools. This was an important moment, as we had also been offered a \$10m venture fund to test some of our hypotheses around strategic project design. After consideration, we decided to turn down the venture fund and focus on training. From June 2014 through May 2015, we ran a number of tests to double-check some of our hypotheses, and by May 2015 had consolidated our efforts.

From that time until the dissolution of the project, our major focus was psychology research, with a special focus on psychologically-informed training. We nevertheless maintained the open research environment, allowing teammates substantial latitude in project selection, on the continuing theory that this was the best way to push research progress forward.

**Phase #3: Self-Application.** From May 2015 through June 2019, we worked hard to make object-level progress on the array of topics that would enable us to run full-stack psychologically-informed training programs. Unsurprisingly, we focused especially on the topics that would let us train people to make progress on developing such training programs. Our efforts resulted in the development of hundreds of models, techniques, and practices, the learning or use of many of which increased a person's ability to improve further.

The apparent success of our research, as well as the continuing optimism and ambition of our team, enabled us to grow to a team of forty-five, raise venture funding for a training- and incubation-based startup, and focus more and more resources on the challenge of human self-improvement.



**Phase #4: Dissolution.** Through 2017 and 2018, our research proceeded methodically, with multiple psychology-focused research teams taking different but complementary approaches. In 2018, we broke open the area of research we now call <u>intention research</u>. This led to increasing divisions within the project, overwhelming our ability to maintain internal coordination despite our serious commitment to understand coordination and help people become better at it. Then, as it looked as though some researchers were suffering negative effects from the intention research, in June 2019 the Executive Director dissolved the project.

The end of the Leverage project was in some ways abrupt. While it had been a plan for some time to move teams in the direction of operational independence, at the end people received one month of notice and in some cases a small amount of bridge funding. After that, most team members went on to different projects. Some team members stayed behind to help reorient and refound the remaining organizations, and over the subsequent months, Leverage was reconstituted. We narrowed our focus, selected new projects, and professionalized. This yielded the institute in its current form, i.e., Leverage 2.0.

For more on the dissolution of our 2011–2019 research project, see Why did Leverage 1.0 dissolve?

### More Information

# **Common questions**

# Why was Leverage 1.0 secret?

Beginning in April 2012, and continuing through June 2019, the Leverage research collaboration operated under a series of information management policies. These policies existed to prevent the misuse of the results of Leverage's research and protect individual privacy, while enabling an open research environment and collaboration between researchers internally.

There was special reason to be concerned about the potential misuse of Leverage's research. In addition to concerns about the risks of psychological experimentation, we had substantial uncertainty around the difficulty of precipitating positive feedback loops of improvement in humans, as well as how powerful such cycles could be. These concerns were amplified by uncertainty around whether there were strategic actors seeking to use new technologies to do harm. Over time, as we gained greater understanding of the risks associated with our research, we diminished the strength of our information management policy.

# Why were people from Leverage 1.0 so optimistic?

In the first few years, Leverage drew optimism from internal and external sources. Internally, we found many promising underexplored research paths, especially pertaining to the possibility of improving the effectiveness of individuals and groups. Externally, we bordered on and hired from the



Effective Altruism community and Rationality community, both of which had sources of optimism of their own.

Over time, Leverage's optimism came to be drawn more and more from internal sources. Our research efforts appeared to be working, individuals testing our psychologically-informed training methods appeared to be improving, and expectations about positive feedback loops seemed to be being realized.

### Why did Leverage 1.0 dissolve?

The reasons for the failure of a project are often complex and difficult to discern, and our answer here may change upon further reflection. Thus far, we have identified the following three key contributors:

- (1) Organizational dysfunction. From the beginning, Leverage had uneven leadership. Our founder was able to find compelling narratives, drive research progress, recruit, and raise funds, but had little skill or experience with management. Team leaders contributed substantially, but were themselves frequently inexperienced. Leverage's decentralized structure was partially a solution to this problem, but ultimately facilitated factionalization.
- (2) Breakdown of internal coordination. Part of the original design of Leverage involved having individuals or groups working on different projects, but in sufficient proximity to and functional communication with each other to yield mutual benefit. The hope was that efforts would then converge over time. This blueprint for long-run consensus and coordination worked for many years, though eventually came under pressure as the organization scaled, and teams developed their own visions and operational capacities.
- (3) Difficulty of intention research. Throughout the project, coordination was an explicit research focus and one of the places we sought to take advantage of positive feedback loops. Through 2017, these efforts were fruitful, and we did not anticipate serious disruption. In 2018, however, we began what we now call intention research, which soon began to yield a surprising and unanticipated discoordination effect. This combined with the potentially negative effects researchers were experiencing from the intention research led the situation to be unsustainable.

As we understand it, these factors were each counterfactually relevant for the failure of the project. Had we been better or more experienced at leadership and management, had teams been more able to benefit from one another despite conflict, and had we avoided or approached intention research very differently, the project may have continued.



### What are the major takeaways from Leverage 1.0?

The following are a few takeaways from our experience with the Leverage project. This is not an exhaustive list and may be something we update in time as more content from the pre-2019 research collaboration is made public. If former collaborators and people who interacted with Leverage 1.0 have feedback on things missing from this section, we would love to hear them. Feedback and questions can be submitted <a href="https://example.com/here-project.">here-project. This is not an exhaustive list and may be something we update in time as more content from the pre-2019 research collaboration is made public. If former collaborators and people who interacted with Leverage 1.0 have feedback on things missing from this section, we would love to hear them. Feedback and questions can be submitted <a href="https://example.com/here-project.">here-project.</a> contact@leverageresearch.org.

- **Explicit problem-solving.** In the course of the Leverage project, we encountered and solved hundreds of problems, including quite tricky problems, by explicitly stating the problem, describing what was hard about it, understanding why others had not solved it (or if they had!), and arguing about which solutions could or could not work. This gave us a lasting sense of self-efficacy with respect to solving problems through careful thought.
- Breaking things down. In many cases, people seem to in some sense believe that things run on magic. We found, both in our research and our practice, that it was possible to figure out how things work, break large tasks into smaller tasks, or big questions into smaller ones. We had so many experiences of explaining something that seemed inexplicable or analyzing something that seemed unanalyzable that we now have a long-lasting optimism about the potential to take steps forward, even when that initially seems impossible.
- Difficulty and tractability of self-improvement. Our attempts to cause a self-improvement cycle with respect to human capabilities led us deep into the weeds of human dysfunction. Our preliminary report, which we believe will be borne out by future research, is that people's minds contain much more dysfunction than one might naively think, and yet are also much more capable of improvement. Overall, we now believe the topic to be broadly safe to explore, in part because it is so difficult, and still extremely valuable.
- **Difficulty of coordination.** Despite the effort, thoughtfulness, and will we put into solving problems of human coordination, we certainly failed, despite working with fewer than fifty people. We take this as a testament to the difficulty of coordination, as well as our own lack of skill, though despite these, we retain optimism here as well.
- Need to continue learning. It may go without saying, but one thing we learned from the Leverage project is how much more there still is to learn. The project was certainly hubristic, in a sense, and we were laid low. But hope springs eternal, and there are problems to which humanity still needs answers. It is thus our hope that others join in continuing research on key topics, though in a form more suited to victory.

The preceding are broad lessons, rather than specific things learned from individual mistakes. For lessons learned from specific mistakes, see the section on Mistakes below.



# What were the major mistakes made during Leverage 1.0?

While in general we think very highly of our early research collaboration, it was imperfect in many regards, including important regards that diminished our effectiveness, caused a breakdown of coordination on the project, and led to negative experiences for some of our collaborators.

Some of the mistakes from Leverage 1.0 were identified through the <u>Experiences Inquiry in 2021</u>, others were discovered through our own reflection. In some cases, conversations with external parties have been helpful as well. Our thanks to everyone who participated in the inquiry or other relevant discussions; it is by identifying where we go wrong that we expect to get better.

Below we list some of the mistakes we have identified, along with their source, an example that typifies the failure, the lesson we learned, and actions taken in response. This list is incomplete; we expect to add to it in the future as we have the opportunity for further reflection and discussion on relevant topics.

### Institutional Design

- Lack of specific vision for the organization see p. 24 of <u>Experiences Inquiry report</u>
- Insufficient central organization and centralized support see p. 25 of <u>Experiences Inquiry</u> report

#### **Institutional Operation**

• Lack of public engagement — see p. 28 of Experiences Inquiry report

### Research Program Execution

- Uneven documentation.
  - Problem Some of Leverage 1.0's actions, research, initiatives, and programs were carefully documented. Much, however, was not documented, leaving some research difficult to communicate and important facts about the institute's history challenging to reconstruct. This error became apparent after the decision to begin publishing Leverage's old research and increasing public communication about the institute's history circa 2019.
  - Example The Leverage research report on <u>Argument Mapping</u> had to be reconstructed from memories, interviews, and a review of more than forty documents. This involved many hours of work, and included many hours of digging through email and document archives to reconstruct the timing and order of key updates and experiments. Had we thoroughly documented the relevant research or prepared it for publication at the time, it would have been much easier to write the research report.



- Solution The institute has a partial solution to this problem, which is to include among its priorities the publication of research and information about the institute. This in turn incentivizes writing up such material, which involves documentation. This solution does not cover the challenging question of how to know which things are worth documenting in advance, or weighing the benefits of documentation against the costs.
- Breakdown in research communication.
  - Problem As Leverage 1.0 did not have any official reporting system for research, in some cases researchers did research that they did not communicate internally in a way that yielded an institutional record. This problem became clear in early 2019 when, as part of the breakdown in coordination, several researchers failed to communicate some of their research to the Executive Director or members of other research teams.
  - Example One researcher did research on the history of psychology; this could have been useful for a number of objectives and would be useful for the institute's work today. However, this research was not communicated, and thus the institute still does not have access to it. Had we ensured continued communication about research, the institute would be in a better position to write and publish on the topic of the history of psychology.
  - Solution This problem is diminishing over time, as some former researchers are communicating their previous research to present institute staff and others are publishing independently. This is of course only a partial solution. Going forward, research conducted as part of the institute's main programs is intended for publication, which means that it should be shared internally by default.

#### Internal Culture

- Disdain for conventional wisdom see p. 23 of Experiences Inquiry report
- Too high tolerance for conflict see p. 27 of Experiences Inquiry report
- Underweighting personal fit see p. 28 of Experiences Inquiry report
- Overfocus on psychological solutions see p. 29 of Experiences Inquiry report

The foregoing are, again, a selection of mistakes made at Leverage 1.0; we intend to expand the list in the future. If there are important mistakes the institute has not yet covered, please feel free to bring them to our attention using the institute's <u>feedback form</u>.



### What are the major controversies pertaining to Leverage 1.0?

There are broadly three classes of controversies relating to Leverage's early work. The first pertains to **conflict with the Effective Altruism community**, the second pertains to **conflict with LessWrong Rationalists**, and the third pertains to **negative experiences** had by some members of the project, especially during the period 2018–2019.

### Conflict with the Effective Altruism Community

With respect to the Effective Altruism (EA) community, conflict arose primarily around competition over talent. Leverage Research was an early contributor to EA movement building efforts, partially for the sake of directly causing positive impact, partially to help the EA movement grow for the benefit of EA, and partially with the aim of finding talented hires. This last point led to substantial conflict, as EA leaders encountered challenges hiring the people they wanted, in a few cases directly as a result of competition with Leverage.

One important example of this conflict centered around the Pareto Fellowship, a summer program run by the Centre for Effective Altruism. Leverage and its former sister organization, Paradigm, contributed staff time and organizational resources to Pareto, in part as a result of perceived shared aims pertaining to the growth of EA and in part for the purposes of finding talented people to hire. From Leverage's perspective, this effort was very successful, with participants reporting very positive experiences with the program and Leverage hiring 4 of the 18 Pareto Fellows as a result of the Pareto Fellowship.

### Conflict with LessWrong Rationalists

Regarding LessWrong Rationalists, conflict occurred primarily around competition over funding sources. Leverage raised \$8,000 from one donor to the Machine Intelligence Research Institute (then the Singularity Institute for Artificial Intelligence) in 2013, and also received funding from the Berkeley Existential Risk Initiative and the Survival and Flourishing Fund in 2018-2020. The amounts raised were small relative to the capacity of the donors and the budgets of relevant Rationalist organizations; nevertheless, Leverage's fundraising efforts contributed to substantial anxiety among LessWrong Rationalists.

An example hotspot occurred in 2021, after Leverage had received a grant from Tyler Cowen's Emergent Ventures for the institute's Bottlenecks work. After word of this grant was made public, and while Leverage was applying for a grant from the Survival and Flourishing Fund, Leverage was subject to a variety of reputational attacks. One of these was by an anonymous LessWrong member who cited the Emergent Ventures grant as the reason for their post, and one was by Oliver Habryka, who was centrally managing the Survival and Flourishing Fund and who in private conversation communicated his desire that people he considered to be Rationalist funders (like Tyler Cowen) not fund Leverage Research.



### Negative Experiences on the Project

With respect to negative experiences on the project, our <u>2021 Experiences Inquiry</u> concluded that of the members of Leverage's 2011-2019 research collaboration, three to five of the approximately forty-five people on the project had very negative experiences. Two of these people have written extensively about their experiences: Cathleen (<u>In Defense of Attempting Hard Things</u>) and Zoe (<u>My Experience with Leverage Research</u>).

In very broad strokes, some of the major sources of negative experiences on the project and after it ended were disappointment from the collapse of the project, internal conflict and tensions rose near its end, and negative experiences associated with <u>intention research</u>. For more about the factors that contributed to the negative experiences and associated mistakes, we recommend reading the full report from the 2021 Experiences Inquiry.

# Culture, Values, and Policies

# What was it like to work at Leverage Research pre-2019?

Leverage Research had a unique internal culture with few formal structures or hierarchies and where researchers had the freedom to choose their own research topics and methodologies with little or no pressure to publish or otherwise distribute the results of their research. For some, Leverage was a thrilling and invigorating intellectual environment and working at Leverage was a strongly net positive experience overall. For others, Leverage could be disorienting and difficult, involving long hours and little separation between work and life to fall back on when work became challenging.

As of August 2022, three participants in the Leverage research collaboration have written substantive accounts of their experience, in reverse order of publication:

- Cathleen <u>In Defense of Attempting Hard Things</u>, the most extensive and accurate account of Leverage Research published thus far, covering 2012–2019
- Jonathan <u>Leverage Research: Context, Analysis, and Takeaway</u>, another account of Leverage Research, covering 2017–2019 and extending into 2020
- Zoe <u>My Experience with Leverage Research</u>, the initial account which sparked the Experiences Inquiry, covering 2017–2019; while we believe the author had genuine negative experiences, the institute disputes this account on much of the substance and implication

Additionally, after Zoe's account indicated that she suffered substantial psychological effects from her participation in Leverage's psychology research program between 2017 and mid-2019, we conducted an inquiry to learn more about the experience of other researchers and identify factors and mistakes that contributed to negative experience. The inquiry found that while people's experience of Leverage 1.0 was generally positive, an estimated three to five individuals had very negative experiences. The inquiry identified nine factors and seven mistakes that contributed to



people's negative experiences; these are detailed in the <u>inquiry report</u>, which is now available online. Our attempts to learn from the mistakes we identified are ongoing.

Following our reorganization in 2019, Leverage has adopted standard organizational practices and structures with much clearer boundaries between work and life. During the pandemic we operated as a remote team, and now have decided to move to a hybrid setup, with some team members working at an in-person office (yet to be acquired) and others working remotely. Early in 2022, we also adopted Google 20% time.