Factors and Mistakes that Contributed to a Range of Negative Experiences on Our 2011-2019 Research Collaboration

Public Report on Inquiry Findings

Executive Summary

In October 2021, a former team member, Zoe Curzi, wrote about her experiences as part of the research collaboration we led from 2011 to 2019, now often referred to as "Leverage 1.0."[1] In order to better understand the circumstances she described, we launched an internal inquiry to investigate the extent and causes of negative experiences during our previous research collaboration.

This report presents the findings of the inquiry, the lessons we have learned from it, and our reflections on some of the important mistakes we made as an organization during that time. As an institute, we believe it is important to understand and take public responsibility for our mistakes both so we can learn from the past, and so others can help us do better by holding us accountable.

Overall findings. The inquiry, led by Paradigm board member Matthew Fallshaw, contacted 45 former and current team members and interviewed 15, having several extensive discussions with many of them. Of those interviewed, almost all reported positive or strongly positive overall experiences. Many described having learned and grown a lot in an intellectually diverse and stimulating environment, developing valuable relationships, and having future opportunities as a result of their participation.

However, most of those interviewed noted that while their overall experience was positive and they would still choose to do it again, their time with Leverage 1.0 nonetheless contained some very difficult aspects and bad periods. Importantly, a small minority reported being unsure whether they would describe their time at Leverage 1.0 as negative or positive overall, having experienced many strong positives but also important negatives and ongoing difficulties they were still processing. It was also particularly significant to us that they reported having worked with other individuals whom they believe may have had overall negative experiences. Further, although interviewees disagreed with the overall characterization of Leverage 1.0 presented in Zoe's account, most believed she was trying to honestly convey her own experience and that others had had experiences that were in some ways akin to what Zoe described.

Investigating people's experiences of Leverage 1.0 further, we came to believe that approximately three to five of the around 43 members of Leverage 1.0 in question may have had very negative experiences overall, particularly towards the end of the project (around 2017 to 2019).

Having established an overall picture, we decided to investigate the problems and negative experiences that had come up, using interviews with former staff, organizational records, and details in Zoe's—and later Cathleen's—public accounts to build a picture of the causes that contributed to negative experiences.[2]

---

[1] External links are marked with an external link icon. All other links are internal links to sections of this document.
[2] Cathleen's piece is a detailed account published in December 2021 by someone who was part of Leverage 1.0 from 2012 through 2019.

Factors that contributed to negative experiences. The picture that emerged was not of any leading cause of suffering but rather a combination of factors that contributed to the risk of negative experiences, including organizational failings and social factors (1-6) and circumstances that added significant stress and conflict (7-9).

1. Lack of Management Skill and Experience
2. Personal Relationships Amongst Leadership
3. Insufficient Institutional Support for Continued Active Choice
4. Pressures on Trainers and Trainees
5. Not Having an Active, Independent, and Legible Support Function
6. Factors that Increase Personal Stakes
7. Intention Research
8. Funding Crunch in 2019
9. Negative Sentiment from Nearby Communities

Our mistakes. Once we received the inquiry's findings, both about people's experiences and the factors that most contributed to negative experiences, we spent time reflecting on what we could learn from it. We identified several mistakes we believe we made during Leverage 1.0, which either contributed to the factors identified by the inquiry or to negative experiences more generally. These mistakes were:

1. Disdain for Conventional Wisdom
2. Lack of Specific Vision for the Organization
3. Insufficient Central Organization and Centralized Support
4. Too High Tolerance for Conflict
5. Underweighting Personal Fit
6. Lack Of Public Engagement
7. Overfocus on Psychological Solutions

Some of these errors had been identified previously and importantly informed the restructuring of the institute in 2019. The inquiry, however, gave us additional clarity on these mistakes and also led us to identify new ones, especially the over-tolerance of conflict and underweighting of personal fit.

Limitations of the inquiry. The report that follows is limited in important ways. Although we reached out to everyone from Leverage 1.0, not everyone chose to speak to us, including those who may have had the worst experiences. Consequently, we are sure that this report is incomplete and misses important elements. We also focus on the causes and lessons pertaining specifically to negative experiences. Therefore, other important topics, like broader lessons from Leverage 1.0, risks of psychology research, and causes of the project's dissolution, are not covered, and there are many aspects of Leverage 1.0 still to share.

Conclusion. Many of the negative experiences at Leverage 1.0 arose from the challenges of being part of a complex, emotionally intense, and unusual project with difficult social elements, as well as the challenges inherent to psychological self-improvement. The inquiry helped us to understand in far more detail the errors we had identified previously and helped us to identify new ones. Thus, while this was a sad topic to investigate, we found the inquiry highly valuable. We expect to continue reflecting on its lessons in the future and hope we will later have the opportunity to speak more with our former collaborators about their experiences at Leverage 1.0—both the good and the bad.

If you have any questions about the contents of this report, please contact Larissa Hesketh-Rowe (larissa@leverageresearch.org) or Matt Fallshaw (leverage-inquiry@fallshaw.me).

Contents:

Introduction

This report shares findings from our internal inquiry into the experiences of former staff who were involved in a research collaboration led by the institute from 2011 to 2019 ("Leverage 1.0") and our reflections on some of the important mistakes we made as an organization during that time.

The inquiry was conducted by Matthew Fallshaw on our behalf, beginning in October 2021, and involved interviews with a third of those we contacted. Matt then conveyed the results of his inquiry to a staff member at Leverage Research, Larissa Hesketh-Rowe. This report was compiled by Larissa, using extensive discussions with both Matt Fallshaw (to cover inquiry findings) and our Executive Director Geoff Anders (to discuss the mistakes and lessons for the institute that could be learned from the inquiry). Both Matt and Geoff fully endorse this resultant report. We acknowledge that the indirect nature of this information chain, while necessary for protecting the confidentiality of those interviewed, may introduce unintentional errors or omissions into the report, in which case Leverage Research will endeavor to provide updates as we learn more.

We encourage readers with less context on Leverage Research to read the background section below to learn about the circumstances surrounding the inquiry. Those most familiar with our history, including former team members and current supporters, may wish to skip to the discussion of the inquiry, its findings, and the discussion of our mistakes.

Background

Reasons for an Inquiry on Negative Experiences

On October 12, 2021, a former employee of Paradigm Academy (a for-profit organization previously part of the Leverage 1.0 research collaboration), Zoe Curzi, posted My Experience with Leverage Research which described her experiences as part of the collaboration from when she joined in 2017 to when the project was disbanded in June 2019. Her post raised concerns about the psychological research, psychologically-informed training, and work environment she had been a part of and suggested that information management policies and concerns about retaliation from Leverage leadership made it difficult for people from Leverage 1.0 to speak up about their experiences.

We were surprised and deeply saddened by the experiences related in Zoe's account. We acknowledge a responsibility to our staff (past and present), supporters, collaborators, and the public, and thus felt it was important for us to better understand the experiences of those involved in Leverage 1.0, to learn what lessons we could, and to take appropriate action on that basis. Such lessons felt especially pertinent given that we had begun preparing a public release of some of the psychology tools developed during Leverage 1.0 (as part of our current Exploratory Psychology Program), and we were thinking about the risks and dangers associated with psychology research as part of that.

The institute chose an inquiry run by an external but trusted party to increase the comfort respondents would feel in coming forward while reducing bias. As part of the inquiry, it also seemed prudent to check

in on current staff to ensure that none were having negative experiences as part of continued training with Paradigm, even though no concerns about that had been raised.

Leverage Research Pre-2019 (Leverage 1.0) Versus Today

From 2011 to 2019, Leverage Research led a sprawling research collaboration investigating a wide array of areas in the social sciences. The environment was informal and unstructured, and researchers were afforded a great degree of freedom in the selection of their projects. Over time, the collaboration came to include many projects, several organizations, and at its peak, almost 45 people. This research collaboration was often informally referred to as "Leverage" or "the ecosystem," and more recently—given the need to distinguish it from Leverage Research today—"Leverage 1.0."

In June 2019, the Leverage 1.0 research collaboration was disbanded, with the various projects and organizations involved instead coming to operate independently and many of the staff who had been involved being let go. Leverage Research, then with only a small number of staff, completely restructured to focus on studying and supporting early stage science and applying the lessons of its research to aid scientific advance. Today, Leverage Research has a more standard work environment with three main programs—History of Science, Exploratory Psychology, and Bottlenecks in Science and Technology—and all of our staff have defined roles within the organization.

This restructure was designed to take into account the lessons we had learned from Leverage 1.0 and address some of the issues that led to dissolution. These included moving towards a more centralized and professionalized work culture with delineated roles and responsibilities, clear reporting lines, set work hours, and standard practices for setting expectations and managing performance. It also involved a narrower organizational focus, separation between work and personal lives, and much greater public communication to help others to understand and engage with our work. Some of the issues that we addressed in the restructure surfaced in the inquiry as causes of negative experiences, though the inquiry identified several additional factors and mistakes as well.

Leverage 1.0: Self-Improvement and the Psychology Research Environment

As references to our psychology research, psychological self-improvement, and psychologically informed training are made throughout this report, some readers may benefit from additional context on these concepts and aspects of Leverage 1.0. For a high-level overview of our past psychology research, see the history section of our Exploratory Psychology Program page.

By 2015 and lasting until mid-2019, psychological self-improvement and psychologically informed training had developed into a central focus within the collaboration. Our psychology researchers worked with hundreds of people, experimenting with a wide range of self-improvement modalities and interventions, including but not limited to testing ideas and techniques with each other or through self-experimentation. Researchers worked to provide assistance on topics like motivation at work, anxiety about romantic relationships, overcoming doubts about personal worth, developing plans for startups, thinking more creatively in research, and so on. This work provided an environment for developing and testing theories about how different people's minds worked.

At Leverage 1.0, researchers were called "researchers" and "trainers." If they were not conducting psychology research themselves, staff engaging in psychological self-improvement were called "trainees." There was a lot of internal excitement around the research, and we aimed to hire people interested in self-experimentation and self-improvement. Our researchers imported methods and approaches from many research traditions and developed many of their own. This meant there was a constant supply of new self-improvement techniques, methods, and interventions for the team to experiment with. While participation in psychological experimentation was intended to be optional, psychology was one of our main research areas and thus one of the main ways staff could contribute to our work. Both Zoe's account and information from those we interviewed suggest that some researchers interested in psychology felt pressure to participate. We discuss some of the problems and pressures of this environment in the section on Pressures on Trainers and Trainees.

Public Accounts of Individuals' Experiences at Leverage 1.0

At the time of writing (April 2022), three former colleagues have written substantive accounts of their own time at Leverage 1.0.

Published	Author	Years Covered	Title
Oct 2021	Zoe	2017-2019	My Experience with Leverage Research
Dec 2021	Jonathan	2017-2020[3]	Leverage Research: Context, Analysis, and Takeaway
Dec 2021	Cathleen	2012-2019	In Defense of Attempting Hard Things

These accounts are currently the best first-hand sources of information about our pre-2019 research collaboration we are aware of. Although interviewees disagreed with important aspects of Zoe's account, we believe it does reflect truths about her experience. As Jonathan and Cathleen's accounts came out later, interviewees had less opportunity to comment on them. We encourage everyone interested in Leverage 1.0 to read these accounts and ask that they be considerate of our former teammates and respect their preferences for privacy, attribution, and how to discuss their writing.

Beyond the accounts above, a few individuals have shared their experiences in places such as community forum comments and an anonymous feedback account set up by someone who knows Zoe. The authenticity of many of these comments is hard to verify but nonetheless might provide useful additional context and insights for those interested. Three comments whose authenticity we have verified and which we recommend reading are those by prevlev-anon, polyphony, and OlliPayne.

We are grateful for the hard work people have put in to share their experiences and help people better understand Leverage 1.0. We hope that, over time, more of our former colleagues will write accounts to help further aspects of the project be understood from new perspectives.

[3] After Leverage 1.0 ended and Leverage Research restructured, we ran a research fellows program. Jonathan, who had been a researcher at Leverage 1.0, stayed on for our 2019-2020 cohort and so stayed with us beyond the Leverage 1.0 project disbanding in June 2019.

Further Information

We have discussed this inquiry and public accounts from former team members in various communications, such as this letter from our Executive Director in November 2021, responding to Zoe's post, and several of our newsletters, which are available on our website.

These communications encouraged those involved to share their past experiences, offered additional points of contact if people were worried about speaking up, and offered support to former colleagues in the form of reimbursements for expenditures on therapy.

The Inquiry

Inquiry Process

Within a few days of Zoe's post, the Leverage Research board convened and asked Matt Fallshaw to lead an inquiry on our behalf. As a board member of Paradigm and longtime advisor, Matt Fallshaw had a deep familiarity with Leverage 1.0 and was known by many of its members, yet was separate enough from the institute to be a more objective and independent person with whom current and former staff could speak. Matt also has extensive management and consulting experience, which could inform his advice to us in terms of ways to improve as an organization.

Initiating the Inquiry: Individuals Contacted

As the concerns centered on the psychology research and associated work environment in the pre-2019 research collaboration, the inquiry contacted everyone who either:

1. Had been employed by an organization involved in the pre-2019 Leverage research collaboration for more than six months and was familiar with or participated in the psychology research (i.e., including all researchers, but not the professional operations team), or
2. Currently worked at Leverage Research and had ever received psychology-based training developed by Leverage 1.0 (such as one-on-one coaching from Paradigm).

As a result, Matt emailed 45 individuals (2 current employees, 43 previous employees) to notify them about the inquiry, encourage them to reach out to him (directly or via an anonymous inbox), and clarify that the informal information-sharing arrangement that was made when Leverage 1.0 ended was not intended to stop people from sharing their experiences.[4]

[4] A copy of the email Matt sent can be found here (external link). The email contains a link to a copy of the Ecosystem Dissolution Information Arrangement (external link). along with notes from October 2021 aimed at clarifying the document's intent and encouraging people to share their stories.

Response Rate and Sources

A third of those contacted agreed to be interviewed by Matt Fallshaw, including individuals from the three primary teams involved in the pre-2019 psychology research (see also: Limitations and Disclaimers). One response was submitted through the anonymous inbox provided.

The insights and recommendations provided by this inquiry are primarily based on those interviews and anonymous response and are supplemented by reviewing organizational records and re-reading the publicly available accounts. As Cathleen and Jonathan's accounts were not shared until the inquiry was nearly complete, they were discussed less in the interviews but contain a lot of useful information we have tried to incorporate in this report.

Reporting Process and Confidentiality

The inquiry was conducted with a default policy of strict confidentiality to try and make it feel safer for individuals to come forward, given that Zoe's post had raised concerns that staff might be afraid of retaliation if they spoke up. Few respondents ultimately opted for less than strict confidentiality, although none indicated concern about retaliation by the institute. To preserve the anonymity of the respondents, direct information from the inquiry, such as quotes from interviews, messages sent, and information about individuals, was rarely shared with Leverage Research, meaning staff at Leverage Research do not have direct access to what was said or who was interviewed. Instead, Matt Fallshaw communicated the findings of the inquiry to Larissa Hesketh-Rowe through a series of conversations; they then worked together to produce notes on the findings and discuss possible recommendations. Leverage staff then used the inquiry findings and recommendations to reflect on mistakes we believe were made and to write this report, which has been shared with Matt Fallshaw to ensure it matches his understanding of topics discussed with interviewees.

Scope and Limitations of the Inquiry

Goals and Focus of the Inquiry

The original goals of the inquiry in October were to (1) determine whether there were any immediate risks or time-sensitive concerns (e.g., if any members of Leverage 1.0 were in immediate danger or any of our current staff had problems with training received from Paradigm), then (2) understand the extent to which others had experiences akin to those Zoe described, (3) determine whether a more comprehensive investigation was needed, and finally (4) gather information that might help us learn from the mistakes and successes of the past.

Having gathered the necessary information on the first three goals, the inquiry ended up focusing primarily on understanding the experiences of individuals from Leverage 1.0—particularly where those experiences were negative—so we could identify mistakes and learn what we might have done to improve the experiences of our former team members. This resulted in identifying factors that led to negative experiences and mistakes that led to those factors, which we cover in the remainder of the report under

Factors and Mistakes below.

The inquiry concluded that a further external or internal investigation was not necessary, though we would very much like to learn more about the experiences of Leverage 1.0 of people who were not interviewed.

Limitations and Disclaimers

The inquiry results and our reflections on them are subject to several important limitations, as discussed below.

Not a comprehensive survey. Our inquiry does not represent a complete survey of individuals involved in the pre-2019 research collaboration. Although Matt reached out to everyone, it has been a few years since the end of Leverage 1.0, and it is possible that some contact information was outdated. Individuals from the three main teams involved in the psychology research pre-2019 were interviewed, but some teams were represented less than others. Matt tried to elicit information from interviewees about the possible experiences of others who did not speak with him, but this has obvious limitations. As such, there are undoubtedly perspectives not covered.

Possible biases and selection effects. There was likely some bias in those who chose to be interviewed. The majority (but not all) of those interviewed already knew Matt. Also, given that interviewees thought it likely that three to five individuals had very negative overall experiences and we did not speak to them, it is possible those who had the worst experiences did not trust the inquiry process enough to volunteer to be interviewed.

Focus on lessons related to negative experiences, not an overall account. The inquiry and this report focused on the lessons we could learn from the negative experiences people had working with us pre-2019. As such, we have not focused on identifying or sharing positive experiences, and we are not trying to present a balanced or detailed account of what the Leverage 1.0 research collaboration was like.

Many important topics are not covered. The focus on negative experiences and well-being also means that there are many subjects this report does not cover, including (but not limited to): the degree of success of object-level projects and endeavors, specific research findings, risks and dangers from psychology research, other elements of the experience of being at Leverage 1.0, or reasons for its dissolution.

Incomplete list of mistakes (even regarding staff experiences). We do not expect this to be a complete account of all mistakes made by Leverage Research that contributed to negative experiences. We were unable to speak to everyone involved in Leverage 1.0, and the events being recounted are from a few years ago (and frequently longer, stretching back over a decade). Additionally, different individuals had different perspectives, and the issues investigated are complex. We have tried our best to be thorough in our investigation and reflections and to cover what we believe are the most important factors and mistakes. However, inevitably the report cannot cover everything, there will still be more to learn, and we hope we may be able to open up further dialogues with our former teammates and colleagues to learn more in the future.

Focus on preventable harms rather than minimizing risk. The inquiry focuses on ways the institute might reasonably have reduced the likelihood that individuals would have significantly negative experiences but not on eliminating risk entirely. Rarely is an organization or individual's goal to eliminate risk completely, but instead to make reasonable trade-offs between different goals. Leverage 1.0 tried to hire talented, agentic individuals who accepted some risks involved in their research. However, this does not remove the organization's responsibility to strive to reduce risk nor the value of learning lessons so as to be better equipped for future endeavors.

Inquiry Findings

Individual Experiences of Leverage 1.0

Overall Experiences Of Interviewees

The majority of those interviewed had either a positive or strongly positive overall experience while working at Leverage 1.0. Their accounts included being part of an intellectually rich and stimulating environment, learning knowledge and skills that they could use in future endeavors, developing valuable relationships, and having opportunities due to their involvement. No significant ongoing risks or dangers were identified, and the current Leverage Research staff interviewed gave extremely positive reviews of the Paradigm training they had received and described positive experiences of working at the organization overall.

These positive descriptions do not imply that people's experiences were always easy. Most of those interviewed noted that while their overall experience was positive and they would still choose to do it again, their time with Leverage 1.0 nonetheless contained some very difficult aspects and bad periods. Two common challenges that particularly came through in the interviews were the work being very hard and emotionally challenging—albeit often rewarding—and feeling excluded or negatively judged by the nearby Effective Altruism and Rationalist communities. Those who had been with the organization longer also described how the environment at Leverage 1.0 became more difficult towards the end of the project, beginning around 2017 or 2018, based on different estimates.

Importantly, a small minority of those we interviewed gave a mixed account of their overall experiences. They felt it was too early to tell whether their time at Leverage 1.0 was net positive or negative, having experienced many strong positives but also important negatives and ongoing difficulties they were still processing. This minority also reported having worked directly with individuals whom they believed may have had overall very negative experiences.

Finally, those interviewed did not feel Zoe's online account was an accurate representation of Leverage 1.0, believing it was misleading both on some details and with respect to its overall characterization, especially where it implied that Leverage 1.0 was a toxic cult, with Geoff as its revered leader. Interviewees had nuanced and differing perspectives on their time at Leverage 1.0, and all felt that Geoff was honestly working to do good, but—in contrast to the impression of him as overly controlling or revered—several commented that they felt Geoff failed as a leader in many respects as he controlled too little and tolerated too much disagreement. Despite this, most felt Zoe was trying to accurately describe her

own experiences and that her account pointed to some important ways Leverage 1.0 had ended up having problematic aspects, even if they disagreed about the implied causes.

A Minority of Very Negative Experiences

Looking into the negative experiences further, predominantly by asking interviewees about others' experiences and from reading the online accounts, we came to believe that around three to five—with two being the lower bound and eight the estimated upper limit—of the 43 individuals involved in Leverage's 2011-2019 research collaboration may have had a very negative overall experience. We believe the worst experiences most likely occurred late in the collaboration, given that those who had been involved longer described a decline in conditions towards the end.

We did not speak to anyone who had a very negative overall experience directly, and no specific individuals were identified to Leverage Research, only mentioned in the context of conjecture to Matt. Nonetheless, many of those interviewed felt the three to five estimate was reasonable.

Understanding The Negative Experiences

Having built a general picture of people's overall experiences, understanding the negative experiences became the focus of our inquiry. We wanted to make sure we understood what we could have done to better support individuals and improve their circumstances. This included both the negative experiences had by people whose experience was overall positive and what happened in the few cases where it seemed likely that people's experiences overall were very negative.

It is difficult to comment fully on what the negative experiences were like, given the complexity and individual nature of personal experience, the length of time that has elapsed since Leverage 1.0, and the fact that we and those interviewed are making inferences. Nevertheless, we believe we were able to obtain at least a partial picture, which we try to convey through the discussion of factors and mistakes below. This partial picture is supplemented by public accounts such as those written by Zoe and Cathleen. Zoe's account, in particular, draws attention to the hard and emotionally challenging nature of the work as well as some of the factors we identified. These include pressures in training, challenges with intention research, and our not being more cautious of factors that increase personal stakes, particularly those related to world improvement narratives and social isolation. Cathleen's account is especially detailed and nuanced, covering many aspects of Leverage 1.0. It includes important background information, illustrations of day-to-day life as part of the collaboration, and insightful commentary on the challenges of running such projects. Cathleen's account also draws attention to the negative impacts of our over-tolerance of conflict, lack of specific vision, and lack of centralized organization.

Although not central to the inquiry findings and outside of the scope of this report, it is important to note that we have encountered, and some interviewees discussed, various surprising and sometimes negative side-effects from introspective techniques and self-experimentation. Such side effects included things like physical shaking, partial amnesia or "brain fog," insomnia, sudden headaches, other psychosomatic pains, and other effects. Not everyone experienced these effects, and those interviewed felt they had consented to the research despite these risks. Where individuals had experienced unusual effects, many did not consider

them to have had a negative impact on them personally. However, these would undoubtedly have added to the difficulty and stress of the work environment.

Key Factors Leading to Negative Experiences

No single dominant cause of negative experiences emerged from the inquiry. Instead, we believe that where our former staff did have difficult experiences, whether isolated or prolonged, they resulted from multiple factors working in conjunction.

The inquiry identified nine major factors that seemed to have most contributed to negative experiences in our pre-2019 research collaboration:

1. Lack of Management Skill and Experience
2. Personal Relationships Amongst Leadership
3. Insufficient Institutional Support for Continued Active Choice
4. Pressures on Trainers and Trainees
5. Not Having an Active, Independent, and Legible Support Function
6. Factors that Increase Personal Stakes
7. Intention Research
8. Funding Crunch in 2019
9. Negative Sentiment from Nearby Communities

Below we describe each factor, the negative impacts it had, and touch on its causes, primarily by linking to the sections where we provide in-depth discussions of relevant mistakes. The first six factors (factors 1-6) are organizational failings and social or cultural factors. The remaining three (factors 7-9) are specific circumstances or events that negatively affected individuals.

(1) Lack of Management Skill and Experience

Many of those interviewed cited a cluster of issues that can be summarized as a lack of management skill and experience as a central problem at Leverage and an important cause of negative experiences. Many felt that our Executive Director, Geoff Anders, managed too little or poorly. Furthermore, staff were not hired or promoted to leadership roles based on management skill or experience but rather based on whether they were already leading projects, and there were few structures or expectations that might have made management easier for those who were inexperienced.

These management deficiencies led to a scarcity of authoritative decision-making at Leverage 1.0 and little central function with the ability or mandate to definitively resolve disagreements, intervene in problems, and clearly and effectively direct the culture. This also meant that how policies were explained or understood and the degree of support individuals received could vary across the research collaboration. Many of those interviewed expressed frustration that while Geoff seemed to have some leadership capabilities, he seemed unwilling (or perhaps unable) to resolve conflicts and make important decisions.

This became more of a problem as the group became larger, and Geoff's ability to set the general direction for the group diminished.

Our lack of investment in management skills resulted from a variety of mistakes. Had we paid more attention to common wisdom, not overly focused on psychological solutions, or had a lower tolerance for conflict, we might have invested more in central organization and centralized support and realized the importance of management and concrete leadership much earlier.

While we think there were many failings of Leverage 1.0 in terms of management, we still have a great deal of gratitude and respect for those who stepped up to lead projects and teams. We recognize that this was frequently a demanding and stressful task that was made all the more difficult by the culture of disagreement, individual freedom, and lack of formal or cultural support for management-style structures.

(2) Personal Relationships Amongst Leadership

Several of those interviewed felt that personal (including romantic) relationships between members of leadership generated work-relevant internal conflict between leaders and made leadership less accessible to the wider team. Some also expressed general concerns that those entanglements may have led to biases in decision-making.

We would normally consider our staff's personal lives to be outside of the scope of the institute's work, including for an inquiry like this one, so long as those personal lives do not impinge on a person's ability to carry out their work responsibilities. We also do not think that close relationships between colleagues are intrinsically a problem. There are many examples of successful couples, housemates, and close friends who work together without this posing issues for the work being done (although there are risks that come with greater personal stakes). However, mixing important personal relationships with work increases the risk of conflict and perceived or actual bias that organizations need to use good judgment to manage, especially where oversight and decision-making authority are involved. If leaders cannot achieve the trust of those they are responsible for, that is an important failure and relevant to this report.

It seems plausible that lack of management skill and experience, which might have guided leadership in separating their personal and professional lives and then effectively demonstrating that separation to staff, was an important contributing factor. We also believe that the lack of an active, independent, and legible support function contributed, as that could have given staff a trusted, independent place to take their concerns, even if their trust in leadership was damaged.

In terms of mistakes, the primary contributors were our disdain for conventional wisdom, which would have recommended in favor of more of a work/life separation, a lack of central support, and having too high of a tolerance for conflict, which led us to underweight the degree to which personal conflict and conflicts in the workplace, in general, can be disruptive and distressing.

(3) Insufficient Institutional Support for Continued Active Choice

Based on discussions with interviewees, we came to believe that where individuals did not feel a sense of agency or power to choose in relation to their work, this likely contributed to especially negative experiences.

The majority of those we interviewed felt that they had consented to the Leverage 1.0 work they were involved in and that they actively and repeatedly chose to engage in that work—including emotionally challenging work—because it benefited them or felt rewarding. Interviewees also felt they had alternative options for what to work on and other trainers they could work with if they wanted to make a change.

However, several interviewees suggested that not everyone felt this way and that those they believed may have had the worst experiences were also those who may have felt the least agency and power to influence their situation. Leverage 1.0 ran much more on informal social structures than formal, centralized ones. As such, there were few organizationally mandated functions for having individuals re-opt into the most difficult aspects of their work or for regularly checking in with some form of independent function aimed at supporting staff to make sure they understood their options regarding changing projects, teams, trainers, or other parts of their circumstances. The emotionally difficult nature of the work and the various challenges of the environment were often discussed, but most often informally and not, as one interviewee put it, in a "unified discussion."

This suggests to us that some individuals may have felt they did not have much control over their environment and may have led some people to feel trapped, isolated, or unable to reach out to make changes. We expand on the impact of this in the case of training and self-improvement in the next section and touch on similar themes in discussing the need for an active, independent, and legible support function.

In terms of mistakes that contributed to this factor, our underestimation of the importance of personal fit meant we did not do enough to hire individuals who would most thrive in the particular environment as it was or help people find better opportunities if it became clear that, even with better support, the environment at Leverage was still not a good fit for them. We also did not provide enough centralized support, which could have given individuals the opportunity to actively re-opt into the work they were doing where it was known to be difficult. We also could have done more to institutionally encourage individuals to consider options other than staying at Leverage, particularly given the many factors that increased personal stakes and made it difficult to consider leaving.

(4) Pressures on Trainers and Trainees

One important factor that made the work emotionally demanding for many and contributed to overly negative experiences for a few relates to various pressures people felt to engage in training and self-improvement at Leverage 1.0. From our interviews, it seems many individuals felt some form of pressure to engage in training or self-experimentation in introspection. Whether this pressure was considered positive or negative by the individuals affected depended a lot on the individual circumstances and type of pressure. It seems likely that where individuals experienced many of these types of pressures

combined or particularly strong pressure to improve in ways they did not wholly endorse, they would have had particularly negative experiences with training.

Both trainers and trainees put pressure on themselves to resolve psychological roadblocks to self-improvement and skill development. The positive and intended version of this pressure is akin to how one might relate to a personal trainer at a gym who helps people meet their own self-improvement goals, applying more pressure when needed but also knowing how to help you ease off if they're doing too much. For many of those we interviewed, this pressure was perceived positively. Where work was emotionally demanding or difficult, this was an active choice they made due to the benefits they anticipated gaining from it.

However, when pressure no longer feels like it serves a person's goals or they do not believe they have the ability to change the goals or the trainer, it's easy to see how pressure, rather than feeling encouraging, can instead feel overwhelming or demoralizing.

One way the pressure may have become negative for some relates to the challenges of navigating the social dynamics. As with other aspects of Leverage 1.0 (such as support systems as discussed in the next section), training relationships and team membership operated on a more social basis than any formalized structure. This means that while some individuals felt able to choose their trainer, team, and work focus, others likely found doing so harder and so felt trapped in training or work situations they did not like.

The socially driven nature of the environment also meant there was social pressure in the direction of engaging with training and introspection. Many people were excited about training techniques and making progress, being a skilled trainer was highly respected at Leverage, and making psychological progress was often praised. Therefore, while there were no formal requirements to engage in training, and some did not, there was clearly a social force pushing in that direction.

Other pressures were likely exacerbated by the fact that self-improvement was part of a widely shared plan within Leverage 1.0 for improving the world. Trainers and trainees often hoped to better equip themselves and others to contribute to the important challenges facing the world today by gaining skills and developing self-improvement techniques others could use. As we touch on in the section on factors that increase personal stakes, any time there is a strong ideological or world improvement-related narrative for one's work, organizations need to be cautious about how this can shape individual decision-making. In our case, the fact that many believed the stakes were incredibly high likely added to the existing pressure to succeed at training.

Finally, many of these pressures led not just individuals to push themselves but also trainers to push their trainees in ways that, in some cases, may have gone beyond what was healthy and supportive of growth. Instances of power imbalances, such as when training with someone who was also your team lead, a member of leadership, or otherwise respected within the group, likely only made this problem worse. We expect such pressure to be especially stressful during difficult circumstances such as the 2019 funding crunch, as evidenced by Zoe's account.

It may be easy to see how these pressures might impact individuals in the role of trainee, but there were also negative impacts on people in the role of trainer. For trainers, the world improvement-related, social,

and other pressures were all reasons to try to do better at training, despite training being something that was being developed on the spot as everyone learned. Further, because there were no clearly delimited roles or boundaries between work and personal life, some interviewees described coming to feel responsible for solving all of everyone's problems (personal and professional), with no easy means of setting personal boundaries. Trainers worked incredibly hard, spending countless hours on what, in many cases, felt like a difficult and thankless task.

We do not pretend to have full solutions to the problems described here. Even under ideal circumstances, getting the right amount of pressure in self-improvement work to support growth rather than causing unwanted suffering is extremely difficult, and our trainers have our continuing admiration and gratitude—without them, much of our work would have been impossible. Nonetheless, we could have done more to reduce the incidence and intensity of negative experiences. If we had consulted more conventional wisdom, we might have done more to set boundaries and navigate power imbalances in doing psychological work. If we had prioritized and developed greater management skill and more centralized support systems, we could have identified and steered away from such pressures. If we had been more cautious with respect to personal stakes, we might have better helped people identify realistic alternative options to psychological self-improvement inside or outside of Leverage.

(5) Not Having an Active, Independent, and Legible Support Function

The majority of those we spoke to felt well-supported during their time at Leverage 1.0 and that they had people to turn to when facing difficulties. However, this support was mostly informal and based on social connections internally rather than formally provided by the organization. This lack of formalized support meant that some individuals might have had less support when needed, and we believe that this was likely the case for the small minority who had the worst experiences at Leverage 1.0.

While Leverage had an employee handbook, dedicated Human Resources (HR) contact, and various policies for supporting staff and raising issues, the HR function was not culturally emphasized by leadership at Leverage (and HR as a concept was not seen by staff as especially relevant or useful), meaning there was no truly active and legible support function. Additionally, from our interviews, it does not seem that the HR-style support that was offered was seen as meaningfully independent from the leadership team. With hindsight—particularly when the collaboration reached a larger size—it would have been valuable to have a support function that was understood as providing valuable, employee-focused support, seen as autonomous from leadership, and able to independently affect policies and culture.

We see underinvestment in centralized organization and support as a key mistake from Leverage 1.0. As with the lack of management skill, if there had been meaningful, organizationally provided and encouraged support available, we might have been able to identify many of the problems listed in this report—such as concerns about management, difficulties in training, or high personal stakes—and worked to address them. Similarly, if we had had less disdain for conventional wisdom and not focused too much on psychological solutions to problems that arose, we might have recognized the value an active and independent support function could have provided much earlier.

(6) Factors that Increase Personal Stakes

Through the inquiry, we identified several factors that increased people's personal stakes in Leverage 1.0. Increased stakes can be positive or negative: on the positive side, one may find one's work more meaningful and be more strongly incentivized to succeed. On the negative side, higher personal stakes can lead to more conflict when there are problems and a greater sense of pressure to succeed while also making it harder for people to leave an environment even if they are unhappy.

Because of the positive and negative elements, the advice from our inquiry was not that these factors should be strictly avoided. Instead, our inquiry recommended that organizations exercise caution with respect to these factors, especially when combined. In the presence of multiple factors, organizations should organize additional check-ins with staff, add guide rails, and do more to make employees feel like they have real alternative options both inside and outside of the organization.

Closely related to the risk posed by these factors is how central working at an organization can become in a person's plans for their future and their vision of a positive life for themselves. When the factors described in this section combine, it can be easy for employees to put all of their eggs into one basket, seeing their work as the primary way they might achieve all of their needs. This can increase the likelihood of conflict and the hurt that can arise when plans don't work out as intended.

We could have done many things to more carefully manage the stakes for the people involved. Generally speaking, our disdain for common wisdom and overreliance on psychological solutions led us to be less cautious than we otherwise would have been about personal stakes. Our lack of centralized support for staff made us less likely to notice problems earlier, and our underweighting personal fit made it more likely we would end up with people who felt heavily tied to the organization despite it perhaps not being the best place for them.

The factors we identified that raise personal stakes, and hence to be cautious of combining, are discussed below.

Social Isolation
If the majority of a person's social and support network is made up of people they work with, leaving the organization they work for can feel as though it risks a significant amount of their social interaction and support. In the case of Leverage 1.0, the fact that the organization fostered a rich social environment, that some had moved to a new city (or even country) to join the team, and the negative sentiment from nearby communities made it more likely that people might primarily choose to socialize with teammates, producing a degree of social isolation.

Close Relationships and Cohabitation
Similar to social isolation, having romantic partners, close friends, or housemates who are also teammates, while having many upsides, can also make the possibility of changing jobs more difficult to consider. Even if someone has social connections across a broad range of different social groups, leaving an organization when there are those relationships may threaten too many important pillars in their life. At Leverage 1.0, many staff had romantic partners, housemates, and close friends among their colleagues.

Intellectual Isolation

One of the disadvantages of developing new intellectual views is that this can make it harder to communicate with others who do not share the same views. As the intellectual views of a group become more distinct, it can become easier to default to communicating with those who already agree with you, leading to an echo chamber effect. During Leverage 1.0, we tried very hard to import new views, inviting in visitors, speakers, and practitioners from different intellectual traditions. But since we did little to communicate our own views (in part because of our lack of public communication and failures in interaction with nearby communities), it became harder and harder to find people to interact with the ideas and perspectives we were developing. This produced an important degree of intellectual isolation at Leverage 1.0.

Strong Ideological or World Improvement Narratives

If a person believes their work is incredibly important (e.g., to the well-being of others) or rare (e.g., if much of society does not recognize its value), it can feel to them as though succeeding at their role in their organization is their only option for having a particular kind of meaningful impact. Leaving a workplace with a social mission can generate a sense of guilt, and so employees in non-profits often tolerate worse working environments than they might otherwise. At Leverage 1.0, many staff members saw their work as an important part of tackling urgent, global-scale problems or as rare, given the uniqueness of the psychological tools and intellectual perspectives being developed. As such, it is likely the perceived stakes of the work made considering alternative options more difficult.

Few Positive Examples Of Others Leaving

Broadly speaking, employee retention is difficult and something organizations frequently seek to improve. However, high retention rates can come with a risk worth considering. If there are available examples of people who have left an organization, are still highly regarded, and are visibly doing well, it can be easier for others to consider leaving. On the other hand, if staff rarely leave or those who do are no longer respected, it can be much harder for people to positively envision working somewhere else. Leverage 1.0 tried to put in place policies that made it easier for individuals no longer officially employed to continue to interact positively with former teammates. In practice, though, Leverage 1.0 had a very high retention rate. Few left the project voluntarily, and few were defunded (let go) before dissolution, meaning there were few positive examples of people leaving Leverage 1.0.

Financial Pressure

Budgetary constraints are a common problem at non-profits, and this can negatively impact staff in a number of ways. Lower incomes make it difficult to keep up with the cost of living, particularly in more expensive cities. If employees are unable to save up contingency funds or the organization is unable to offer severance pay, leaving can incur a serious financial risk. While Leverage 1.0 did have a policy of providing three months' salary to those who had been with the organization for more than a year (or one month otherwise), not all of those interviewed recalled this policy, and so it was likely not always effective in reducing financial pressure. Additionally, the funding crunch in 2019 meant we had to suspend this policy except in dire cases where people were encouraged to speak to leadership about their specific circumstances. In these instances, some people asked for and were given a few months' financial support, but nonetheless, it was clear there was a lot of stress for employees around funding at that time.

(7) Intention Research

Beginning around 2018, many of Leverage's psychology researchers began research in an area that the institute came to call "intention research."[5] These investigations focused on interpersonal psychological effects, and in particular on the importance of subtle nonverbal communication channels and how these were influenced by changing one's "intention." As a result of this work, many of Leverage's psychology researchers came to believe that nonverbal communication is at least as expressive and psychologically central as verbal communication and that nonverbal communication can occur without the conscious awareness of either party yet, nonetheless, cause profound positive or negative psychological effects.

The Leverage 1.0 research collaboration was disbanded in the very early stages of this research, partly due to the emotional distress it seemed to be causing some team members. As a result, substantial theoretical disagreement remains about both the nature of the phenomena involved and how best to explain them, making it harder to describe intention-related experiences as part of this report.

Several of those interviewed felt that intention research had led to a significant increase in internal conflict starting in 2018, and some described a clear causal connection between the worsening of the environment towards the end of Leverage 1.0 and the investigation of this new research area. Thus, in broad terms, at least, it seems likely that this research area was related to the communication breakdowns, the higher degree of conflict, and the degree of demoralization experienced at Leverage 1.0 towards the end.

In terms of conflict and breakdowns in communication, learning to pick up on subtle nonverbal communication from others led some researchers to believe they had become able to detect additional information, including negative opinions and intentions, from people around them. The apparent detection of negative information led both to an increase in conflict (when this negative information was strongly felt or discussed) and parties avoiding one another (in order to diminish unintended, problematic communication). Some were convinced of the reality of the phenomena, while others felt that this period involved less rigorous research and that, instead, people got carried away with their concerns about the implications of this research area. Regardless, many agreed there was an increase in conflict and a decline in communication even if they did not agree on the cause.

In terms of demoralization, many of those who had worked at Leverage 1.0 were excited about the progress they had been able to make using various introspective tools they had discovered or developed. The successes to date led many to believe that it might be possible to eventually—albeit with great difficulty—solve the majority of a person's psychological issues and roadblocks. The psychological material apparently detected during intention research instead suggested doing this would be much harder than previously thought. In particular, it suggested that people were much more influenced by each other than previously imagined, such that solving one person's issues might require solving issues on a society-wide level.

Given the aims of Leverage's psychology research and how unusual and surprising some of the intention research phenomena were, this research area and its negative effects may have been difficult to avoid. Nevertheless, we believe we made mistakes that increased rather than ameliorated the difficulties this

[5] While Leverage Research describes this area as "intention research," not all of our former psychology researchers would characterize it this way.

research area caused, such as not having centralized support mechanisms, having too much tolerance for conflict, having too little external engagement (making discussing this research with others harder), and focusing too much on psychological solutions.

As many of those interviewed expressed a strong desire for there to be an accurate public understanding of their past work, we have recently produced an essay, "Introduction to Intention Research" *[forthcoming]*, in the hope that this starts to make this area of research and aspect of Leverage 1.0 easier to talk about. This piece describes intention research in general, some unusual phenomena and negative psychological and psychosomatic effects encountered, ways intention research contributed to internal tension at Leverage 1.0, and risks posed by further research.

(8) Funding Crunch in 2019

In February and March 2019, the delayed receipt of previously pledged funds and delayed efforts to raise further funds, especially due to internal conflict, led Leverage 1.0—then with more than 40 staff—to run out of runway. In response, we accelerated our existing plans to have individual teams raise more of their own funding independently (for instance, through training and consulting). According to our interviews and Zoe's public account, the funding crunch, and our handling of it, served to create additional stress as individuals faced more pressure to provide value or potentially face defunding.

While funding crunches and layoffs are never easy, we believe that if Leverage 1.0 had more management skill and experience among leadership and more central support for staff, we might have been able to communicate better and earlier across the organization, and thus ease some of the stress. If we had understood the factors contributing to the increased personal stakes, we might have done more to reduce and monitor these risks, making it easier for people to transition out if the circumstance became too difficult. Had we better understood the challenges around pressures in training, we could have intervened to help prevent people engaging in psychology research from putting undue pressure on themselves or each other and advised them on how to better navigate the circumstances. Finally, had Leverage Research engaged in more public communication, we could have built a broader base of support rather than relying on a small pool of large donors, thus reducing the risk of sudden funding crunches of this type.

(9) Negative Sentiment from Nearby Communities

Finally, feeling excluded or negatively judged by the surrounding Rationalist and Effective Altruism communities was one of the types of negative experiences that were among the most common for interviewees, and this is something discussed in a section on the topic in Cathleen's account. Many cited their expectations about negative reactions from these communities as the reason they did not want to speak publicly about their time with Leverage 1.0.

These two communities are ones with which Leverage Research has a history, having collaborated with organizations central to each in our earlier years and having some ideological overlap on thinking carefully about how to improve the world and psychological self-improvement. This overlap meant that, to varying degrees, some likely felt part of these communities at some stage in their lives or had friends within them, making negative sentiment from individuals in these communities understandably difficult.

It is important to note that this factor may have been raised more frequently as an important contributor to negative experiences than it would have been otherwise due to recent events pertaining to the relations between the communities and Leverage 1.0. For instance, while the inquiry was being conducted, staff from some of the central Rationalist organizations organized an $85,000 cash bounty for information about Leverage Research, which was broadly interpreted negatively by members of Leverage 1.0. One interviewee also reported recently being denied a job at an Effective Altruist-identifying organization specifically because of their association with Leverage 1.0.

While we believe the negative sentiment from these communities is unjust, we nevertheless believe that we bear some responsibility. Had we engaged in more public communication, there would be less room for false and misleading narratives about us, and had we had a lower tolerance for conflict, we would likely have distanced ourselves from the more toxic aspects of these communities earlier.

Additional Potential Factors

The preceding list of factors are those the inquiry identified as the main causes of negative experiences on the Leverage 1.0 project. Beyond those, it is worth mentioning one additional factor, Leverage 1.0 information management, which, while not identified as central by the inquiry, we nevertheless believe was an important contributor.

Information Management Policies

Information management was mentioned in Zoe's account as a source of distress. Her account and our conversations and reflections on the topic have led us to believe that we made mistakes in this area that are important to understand and share.[6]

Leverage 1.0 had an information management policy (first introduced in 2012 and substantially revised in late 2015/early 2016) that explained what information could be shared externally and the rationale for the policy. When Leverage 1.0 ended in June 2019, some staff also signed an Ecosystem Dissolution Information Arrangement aimed at encouraging cooperativeness and respect between everyone after the emotionally difficult end of the project.[7]

The original information management policy was created because we were uncertain, but concerned, about our potential misuses of psychological information, knowledge, or techniques. Where in doubt, we decided to be cautious, reasoning that overly stringent policies were more reversible than sharing information, only to later discover this had been a mistake.

[6] This public account (external link) about a different organization, which was inspired by Zoe's post, influenced our thinking, primarily by making us more concerned about the ways information management policies can have severe, unintended negative impacts on individuals' mental health and well-being.
[7] A copy of the Ecosystem Dissolution Information Arrangement (along with clarifying footnotes added in October 2021) can be found here (external link).

We understood that a policy that discussed information management principles and potential risks of sharing information might lead people to be overly cautious in sharing information, but we significantly underestimated the strength of this effect and the negative impact it could have. We aimed for a policy that would be simple to understand and easy to implement, but ended up with a policy that was more difficult to implement in practice than we anticipated without just rounding it off to sharing very little. This led people to sometimes simply not share information about their work with others at potentially high personal cost.

One mistake we made in this connection was overestimating the degree to which the policies would be understood and implemented as intended on the basis of the official statements of the policy, and so we did not do enough to monitor or correct how they came to be understood informally.

Perhaps more important, though, was our failure to recognize a reclusive tendency in our culture or to see that tendency as a problem. As a result, we did not pay as much attention as we should have to the way our information policies made it easier for people to cut themselves off from others. Had we understood reclusiveness as more of a problem, we would have proactively crafted our information management policies to counteract rather than feed into this.

We think there is more for us to learn about how to create and manage these kinds of policies well, should information management policies be required for the institute's work in the future. It is often difficult to balance legitimate concerns pertaining to dangers that arise from sharing information with the needs of the people working with that information.

Mistakes

On the basis of the information received from the inquiry, including the factors identified above, key members of Leverage's present staff, including our Executive Director, spent time reflecting on the primary mistakes we had made that contributed to negative experiences at Leverage 1.0.

We ended up identifying seven relevant mistakes and errors:

1. Disdain for Conventional Wisdom
2. Lack of Specific Vision for the Organization
3. Insufficient Central Organization and Centralized Support
4. Too High Tolerance for Conflict
5. Underweighting Personal Fit
6. Lack of Public Engagement
7. Overfocus on Psychological Solutions

Just as there was no central factor that caused most of the negative experiences at Leverage 1.0, there was no single mistake that explained all of the others. Rather, there were a collection of mistakes, many of which fed into each other, that in turn led to the factors identified by the inquiry described above.

Some of these errors we had identified previously and importantly informed the restructuring of the institute in 2019. Through this inquiry, however, we added depth to our existing understanding of some mistakes and identified new ones.

With respect to mistakes previously identified, by late 2019, we had identified disdain for conventional wisdom, a lack of specific vision for the organization, insufficient central organization, and a lack of public engagement as key failings of Leverage 1.0. These led us to adopt a clearer mission and set of programs, establish centralized management and official roles, and begin publishing our research and communicating publicly about our work. The inquiry, however, led us to recognize further lessons and related issues. With respect to vision, we realized that we also need a concrete vision of what we want the organization's values, culture, and social environment to be like and an image of what it looks like to be a good team member. With respect to central organization, we realized that our current staff would benefit from organizational functions devoted specifically to employee-centric support. Finally, we came to better understand the role public engagement can play not just in research but also in building a base of supporters whom the organization can rely on in times of need and who can defend it against negative sentiment.

We also identified two new mistakes. We learned how having too much tolerance for conflict can have a significant negative impact on a work environment, allow spillover from the conflict that can arise interpersonally, and lead to tolerating unhealthy relationships with nearby groups. We also realized that we had neglected the importance of personal fit as part of our hiring and that this must be taken into account if we are to create a culture which all of our teammates can contribute to and in which all of our team members can thrive. Moving forward, we expect these realizations to help us continue to shape the institute, both internally and in terms of how it interacts with the world.

With respect to the over-focus on psychological solutions, this was a conclusion suggested by our intention research and finally reached in early 2020. The inquiry did not substantially impact this conclusion, though it did provide the occasion to state it clearly and communicate it to the public.

Below we describe each of the mistakes and try to draw a lesson from it for the future. It is only by reflecting on our mistakes and learning what we can from them that we can do better as we attempt new, ambitious projects in the future. We also do not pretend to have fully learned the lessons below nor to have identified all the mistakes we have made; rather, we hope we have made some progress already and intend to make more progress over time.

Disdain for Conventional Wisdom

One of the largest and most pervasive mistakes we identified from the inquiry was a disdain for common wisdom. "Standard solutions have failed," it was frequently noted, "so we need to invent something new." To a degree, it is certainly the case that innovation and new solutions are needed. But we took this too far and, as a result, failed to benefit from the implicit wisdom embedded in common practice.

For instance, even though we were studying the mind and self-improvement and studied a wide variety of psychological traditions and practices, we did not focus especially on best practices in therapy or the field

of psychology, and we likely suggested that team members seek external psychological support less frequently than we should have. Or, even though we were trying to learn how to train leaders and become leaders ourselves, we imported very few (if any) practices from existing managerial approaches.

This contributed to several of the factors that led to negative experiences on the project. Had we paid greater attention to pre-existing wisdom, we might have been able to learn to be better leaders and managers and might have been able to identify and reduce undue pressure among our psychology researchers. We also would have taken greater pause before creating a social group whose world-improvement plans, social plans, and viewpoints were so closely related and entangled with one another or before delving so quickly and deeply into intention research; these are things normal organizations tend to avoid. Neglecting common wisdom also contributed to other mistakes, such as providing too little organizational support, avoiding public engagement so completely, or relying as much as we did on the expected efficacy of psychological self-improvement.

We failed to take advantage of common wisdom as much as we could have for a number of reasons. Understandably, we encountered the objective difficulty of knowing which pieces of common practice to adopt and which not while trying to do something new, especially when the justifications for common wisdom are often implicit while we were trying hard to make our reasoning for actions explicit. But we were also a band of rebels and rejects, more negative on society and normal solutions than we should have been, for reasons unique to each of us but with the inevitably common themes.

Doing something as intense and original as we were trying to do called on us to do something new, and hence different, but also called on us to rise above our own distaste for normal solutions, and in the latter, we could have done much better.

Going forward, we as an institute are working hard to incorporate the knowledge and wisdom of common practice, adopting more traditional norms, studying best practices, and creating a recognizably professional organization. We have made strides since 2019 when we started to professionalize. It is very difficult, however, to unify the good sense of common practice with the insights that arise from explicit reason and to know where new approaches are truly needed and where standard practices would provide a better outcome. We expect to be working on this for some time.

Lack of Specific Vision for the Organization

Another mistake we identified was not having a specific enough vision for the organization. We had a very general long-term vision (i.e., universal flourishing), a compelling mid-run vision (i.e., a two-thousand-person team of effective coordinated individuals), and a plan (i.e., acquire knowledge, engage in self-improvement and training). But we did not have much of a specific vision of which projects might be included, what sorts of contributions we expected team members to make, or how we wanted the internal culture to be in practice.

This contributed to great fecundity, with an explosion of ideas and initiatives, but could also be disorienting and demoralizing. New hires would sometimes experience a collapse of motivation after

joining, sometimes for many months, before figuring out what to work on. Others started many abortive projects, or flitted from project to project, unable to find a stable and productive home.

Had we had a more specific vision of what we were trying to do, it might have been much easier to help people design their own projects, navigate the social space of the organization, and develop a sense of being part of a team rather than being a collection of individuals. Had we had a sharper picture of what success for team members looked like on the project, we might have been able to give people much more specific day-to-day guidance. Altogether, the lack of specific vision increased the requirements for individual agency, making the environment much harder to navigate than it could have been otherwise. The lack of specific vision also contributed to other mistakes, such as our avoidance of public engagement, since it was hard to explain something so open-ended, underweighting personal fit, since without a specific vision, it was easy to imagine that anyone could make a place for their style of work in the project, and our having too much tolerance for conflict, since without a more specific vision we felt the need to house a large number of competing viewpoints.

As far as why we didn't have a more specific vision—for which projects to run, for how we should work together as a team, for who would flourish on the project—part of the answer was that we didn't know what the specific vision should be, and we knew we didn't know. This is the virtue of Socratic wisdom, but it also came with a vice, which is that we, and our Executive Director especially, was afraid of being wrong and being judged for being wrong.

Coordination among people, at least people as inexperienced as ourselves, requires more rails than we were brave enough to provide. It is essential to recognize the limits of our knowledge, but it is also necessary to hazard enough guesses to enable us to navigate the world together.

Leverage Research now has a more specific vision of its activities, with legible roles, discrete programs, and so forth. There is still more work to be done, though, especially in defining the company ethos and developing a more specific vision of a desirable future. We have been working together as a team to discuss company values and discussing possible visions of the future with collaborators and expect to continue to do so.

Insufficient Central Organization and Centralized Support

Another mistake was our insufficient investment in central organization and support. This had pervasive effects, contributing to all of the factors we identified in the inquiry and compounding many of the other errors, including lack of specific vision for the organization, too much tolerance for conflict, lack of public engagement, and overfocus on psychological solutions.

In terms of central organization, Leverage 1.0 did have a small, valiant operations team that managed our accounts, legal compliance, buildings, catering, and much more. However, the purpose of this team was not to directly support and manage the researchers but rather to create an environment in which they could make their own choices. Among the researchers, individuals largely decided for themselves what to do. There were few policies and hence few guardrails. Even where there were policies, we often failed to

predict or productively steer the ways in which social dynamics impacted how the policies were informally understood (or misunderstood or forgotten).

There were also many sources of support, such as trainers and facilitators. We even tried adding various functions that focused on collaborator well-being, such as one of the early versions of Coordination Team, or in the later System Vitals.[8] But none of these sources of support were part of reliable organizational functions which it was part of anyone's official responsibility to maintain or interact with.

This meant that when a funding crunch came in 2019, there was no centralized management to make sure individuals understood what this meant for them, or to provide team leads with guidance and support when navigating this with their teams. There were general announcements, and team leaders spoke to their teams, but what they said and how it was managed was left up to them. Similarly, when it came to personal lives, there was little organizational guidance, and nothing that discouraged personal relationships amongst leadership or in general, or the mixing of many factors such as worldly ambitions, ideological goals, and personal lives. In this sort of environment, requirements for individual agency were high, social factors could apply undue pressure (including in psychological research), and if we did not already possess management skill and knowledge, organizational functions did not exist to provide training and support.

We might have been able to prevent some of the negative experiences on the project had we devoted more resources to centralized support. For instance, we could have created a more active and independent support function, perhaps under the purview of our operations team, with the responsibility to regularly check in with people. We could then have encouraged, or even required, team members to attend regular check-ins. This would have made it possible for us to receive more feedback on the environment and make changes more quickly if needed. It would also have made it easier to communicate the few policies we did have and ensure that everyone understood their options and received help if they needed to switch teams, take time off, or leave the project if necessary.

Required regular check-ins might seem like a small change, but it would have represented a significant cultural shift—and this actually gets to the heart of the mistake as we understand it. The open and undefined environment was designed to give people a lot of freedom, and many of the people we hired were squarely allergic to central organization. (People often referred to anything that smacked of central organization as "the system," with attendant distrust.) We thus avoided structure wherever we could. But we now believe that too little structure, even in an environment rich with affordances, can undermine rather than support agency.

Creating free and open research environments, and capturing the benefits that those bring, requires understanding when centralized functions undermine agency and also when such functions support it. Striking the balance between the two is essential if we are to create organizations that allow agentic individuals to develop and flourish.

[8] From 2014-2017, Leverage 1.0 had a team called Coordination Team, which took different forms at different times. Early on, Coordination Team focused on trying to identify what might make specific team members' lives better through small, thoughtful actions. For a few months in 2016-2017, there was a team called System Vitals whose members tried to regularly check in with other people on the team to see how they were doing. This function was later replaced in 2017 with a new version of Training Team.

The institute now has swung in the opposite direction and is grappling with the challenges from the other side. How, with defined roles and responsibilities, quarterly goals, and a more limited scope, will we create the freedom that yields the benefits of an open research environment with agentic individuals? We believe there is a way to unify both sides, though, as with the other lessons, getting this right will take time.

Too High Tolerance for Conflict

Another pervasive mistake was our tolerance for conflict, which in retrospect, was far too great. Some degree of conflict is inevitable, especially on ideologically charged projects that house many competing viewpoints, like Leverage 1.0. However, the degree and types of conflict are usually minimized by many means, such as having superiors who will step in and make decisions to resolve disputes, or institutionally discouraging conflict that is overly emotional or personal.

During Leverage 1.0, we did few of these things, aiming instead to allow people to work out conflicts of many types on their own while at the same time taking the resolution of conflicts as a legitimate object of study. This meant that conflicts could last for a very long time, be emotionally charged or confrontational, and include personal elements that were not sharply distinguished from legitimate disagreements about work. As a result, the Leverage 1.0 environment could be challenging for many. If we had been more cautious with respect to conflict and tolerated it much less, we would have prevented some of the negative experiences people reported and ideally have had a happier and more productive work environment.

Because of our high tolerance for conflict, which was aided by the lack of central organization, we did not avoid personal entanglements, including amongst the leadership team. We also were not as wary as we should have been of factors that increased personal stakes or of the way personal conflict could contribute to problems with pressure in training. Our tolerance for conflict would also likely have added to the emotional challenge of the funding crunch in 2019 and intention research, and led people to put additional weight on psychological self-improvement as a way of solving problems.

We also wonder whether this mistake influenced our approach to handling negative sentiment from and conflict with nearby communities in a way that yielded more negative experiences for our former staff. If we had less tolerance for conflict overall, we might have stood up for ourselves or distanced ourselves from those communities more quickly.

Our over-tolerance for conflict arose partially from insufficient management skill and experience but more centrally from a misunderstanding of the nature of conflict and its effects. Conflict is often useful and sometimes necessary. But it is much more pernicious and destructive than we, and especially our Executive Director, initially believed, especially if left unaddressed for long periods.

Today, with the benefit of hindsight, Leverage Research aims for much less conflict overall. Internally, we have professional norms and decision-makers with the authority to resolve conflicts. Externally, we are working out practices and standards for dealing with conflict in the public domain. In both, we expect to take into account our newfound recognition of the value of conflicts being brought to resolution.

Underweighting Personal Fit

Reflecting on the range of negative experiences people reported led us to realize we had importantly underweighted considerations relating to personal fit and personal style in our recruiting. We had hoped that Leverage 1.0 would be sufficiently freeform and changeable that most preferences could be accommodated, but ultimately the freeform nature of the project itself and our main research focuses shaped the project in ways that gave it a definite culture.

We recognized this to an extent, and this was reflected in our hiring criteria and process where we, for instance, tried to make sure people would be okay with emotionally challenging work and discussed many of the more difficult aspects of the Leverage environment, such as the self-supervised structure, lack of consensus, and modest pay. On reflection, however, this was more geared to checking whether people could survive the culture, not whether or not they would thrive in it. It would have been much better if we had tried to find people who found the environment fit their preferred work style, rather than relying on people to adapt. On the other hand, it would also have been better if we had done more to steer the culture in line with what staff needed, such as more structure and less conflict. Had we done so, we think some of the most negative experiences at Leverage 1.0 could have been avoided.

The effects of underweighting personal fit were likely made worse by the factors that increased personal stakes since high personal involvement could make it hard for people to leave, even if that would be better for them.

Our failure to appropriately weight personal fit was, unfortunately, a natural mistake, given our focus on psychological self-improvement and organizational design. We too quickly assumed that problems people encountered were a result of them or the environment and thus could be overcome by helping the people to change or changing the environment. However, our intention research, in particular, showed us how connected individual development is to the environment and other people, which means that in some cases, the correct solution is to help people find environments they will naturally flourish in, rather than trying to adjust one or the other.

At Leverage Research today, we take account of personal fit in our hiring as best we can. We also have a substantially simpler and better defined organizational culture, which makes judging personal fit easier, and a number of features (including a clear work/life distinction) that make it easier for team members to consider alternatives in case working elsewhere would be better for them.

Lack Of Public Engagement

Although we had already identified our neglect of public engagement as a mistake—leading us to prioritize it since 2019—we hadn't realized the extent of its negative impact until learning more about people's negative experiences through public accounts and the inquiry.

Our original rationale for neglecting public engagement was a cost-benefit analysis. We had tried public communication in our first few years, especially 2011-2013, but a review of this led us to judge it as costly in terms of time with few benefits. We then broadly neglected our public footprint until, in 2019, we realized our mistake and changed plans. Even then, however, we had only really grasped a few general

reasons, such as the importance of public accountability and the value of keeping the public apprised of our findings even when provisional.

The public accounts of experiences with Leverage 1.0 and reports from interviewees made the value of public engagement much more evident and real. The fact that we provided so little public information about the project made it hard for members of Leverage 1.0 to talk about and make sense of their experiences with other people, and also left room for others, including those with whom we had conflict, to create misleading narratives that were difficult to combat. This led some interviewees to report feeling socially isolated, bullied, excluded, or judged by others, which in turn fed negative cycles, with some feeling less comfortable speaking about their experiences and some believing there was less value in doing so.

Our previous lack of public engagement was enabled by our over-tolerance for conflict, and made it so we had only a narrow base of supporters, which was a partial cause of the 2019 funding crunch. Had we stuck with public communication, despite our initial failures, we would have been able to build a large base of support, with many people understanding our work and supporting our team. This would have reduced social isolation, made it easier to push back on misleading narratives about us (or prevented them in the first place), and provided backup funding sources.

We think we avoided public engagement for a combination of reasons. In addition to the initial cost-benefit analysis, we also found that as Leverage grew, it became possible to get high-quality engagement internally without having to brave the untamed wilds of the internet. At the same time, we also severely underestimated exactly how negative external sentiment could become, how long it would last after we stopped engaging with the relevant groups, or how that might affect some of our team; had we understood this, we would have taken a very different approach from the beginning.

We have taken some steps to communicate about our work, past and present, to the public. This includes our regular newsletter, the write-ups of some of our past and present research, and others sources of information such as our Twitter account. We have also started to build a larger base of support, as evidenced by our recent first successful public fundraiser. Others have helped, especially those who posted public accounts of their experience at Leverage 1.0. Future steps are planned as well, including the dissemination of our introspection tools and more publicly available information about Leverage 1.0 and Leverage 2.0 (i.e., the institute today). We expect it to take substantial further effort to communicate the relevant information, both on the part of ourselves and of others, but through those efforts, we believe that in time the entire story of Leverage 1.0 will be told.

Overfocus on Psychological Solutions

Finally, one central mistake we made was to focus too much on psychological solutions to problems. Our research gave us very good reason to believe that many problems — emotional, attitudinal, and motivational — have psychological solutions. It also gave us reason to believe that many skills could be built and desirable character traits developed through psychological self-improvement.

However, we were substantially less good at solving problems in other ways, and this led psychological solutions to become the go-to approach when problems arose. When dealing with insecurity about one's work, or concerns about the future, it was easy for individuals to identify irrational beliefs and attitudes at play for them and take steps to overcome them. This approach became more and more viable the more we improved our psychological interventions.

As it became commonplace to solve problems by building skills (via psychological means) or overcoming unendorsed viewpoints (via psychological means), it became natural for people to try to adjust themselves to undesirable circumstances using psychological self-improvement. This was such a strong attractor that it became frequent advice that trying to debug oneself out of having one's own desires was a fool's errand (a "rookie move," it was often called) and almost certainly harmful. But this did not stop people from trying.

The affordance for psychological self-improvement, and the benefits the tool had in many domains, contributed to some of the factors that led to negative experiences on the project. This was by different means. It contributed directly to undue pressures in psychological research, where the availability of psychological solutions and a variety of challenges and reasons to improve led people to see psychological self-improvement as a natural option and thus not adequately consider alternatives. It contributed indirectly to underinvestment in other solution-types and hence to insufficient management skill and experience; had we not relied so heavily on psychological self-improvement, we would have been forced to develop other skills. It also led indirectly to us being less cautious with respect to factors that increased personal stakes, or in investigating intention phenomena as we assumed that problems with either could be managed at least in part through introspection and self-improvement. With respect to intention research, as well, our reliance on psychological self-improvement rendered us vulnerable to demoralization, especially when intention research raised the possibility of a vast and practically insurmountable number of psychological issues to address. The affordance for psychological self-improvement also contributed to or compounded some of our other mistakes, including disdain for common wisdom, over-tolerance of conflict, and underweighting personal fit.

The core of the mistake here was a factual one. We underestimated the depth of human psychological complexity and dysfunction, largely because our initial estimates of the difficulty of psychological self-improvement were based on our maps of the conscious mind, rather than the conscious and unconscious mind.

Going forward, the institute no longer has psychological self-improvement as the centerpiece of its strategy. We believe that psychological self-improvement techniques and training are highly valuable and find it beneficial to have these available (optionally) for our employees by external providers. But we also now believe that solutions to humanity's problems will come from a broader range of sources, with psychological self-improvement playing a small, though not insubstantial, role.

Conclusion

This report has covered the results of our inquiry into negative experiences had by members of the research collaboration spearheaded by Leverage Research ("Leverage 1.0") from 2011 to 2019. Many of the negative experiences at Leverage 1.0 arose from the challenges of being part of a complex, emotionally

intense project with importantly difficult social elements, as well as the challenges inherent to psychological self-improvement.

We found this inquiry extremely valuable. Prior to the publication of Zoe's essay and this inquiry, we knew that some staff had had a difficult time prior to and after the dissolution of Leverage 1.0, and had reflected on mistakes we had made as an organization during that era as part of thinking about how we could build a better and more sustainable organization. The inquiry, however, as well as the public accounts, gave us a much more complete understanding of both the causes and extent of negative experiences on the project.

The current form of the institute—Leverage 2.0, as it is sometimes called—embodies the lessons that we had learned after the collapse of the Leverage ecosystem, including some of the mistakes described in this report. We have tried to build an organization with a more specific vision, with much more structure and support for staff, that engages with the public, and that pays greater heed to conventional wisdom. We had also come to, through reflection on our research, the conclusion that psychological solutions should play a much smaller role than we initially had hoped in 2011-2019.

The inquiry underscored the importance of these lessons, not merely abstractly, but in terms of how mistakes on these dimensions led people to have worse experiences on the project than they would have otherwise. It also brought into focus two further errors: our underweighting of personal fit, especially in the context of high personal stakes, and our over-tolerance for conflict. Going forward, we expect these realizations to impact the institute, both in terms of internal culture and in terms of how we engage with external groups.

The topic of this report—the extent and causes of negative experiences at Leverage 1.0—was obviously a sad one to investigate. Beyond informing our future endeavors, we hope that the lessons described here will also be useful to others, especially individuals and groups running ambitious projects in unstructured domains. We also note again that we were not able to speak to everyone from Leverage 1.0, and so we hope to learn more about our former teammates' experiences, both the good and the bad, as conversations continue into the future.

Acknowledgments

We would like to thank everyone who contributed to this report, including everyone who has shared their experiences publicly, all of the interviewees who spent many hours in conversation discussing their experiences and surrounding context, and especially Matt Fallshaw for conducting the inquiry on the institute's behalf and investing so much time in helping us learn the lessons we could from it.