

Worldview Investigations and Data and Surveys Analysis Update

Quarter 3, 2024

Broadening our moral circles

A key motivation for Rethink Priorities in launching the [Moral Weight Project](#) was the recognition that people tended to prioritize helping humans or other animals based on *intuitions* rather than *evidence*.

Using a multidisciplinary approach, **the research team found the estimated welfare capacities of the species they examined to be relatively similar to one another.** By acknowledging the challenges inherent to such comparisons, the researchers sought to provide a transparent foundation for prioritization. RP also produced graphics to make the findings more accessible. Thanks in part to staff outreach, **the project's findings gained attention among animal advocates and generated discussions** within the field at-large.

- The research has been published in academic venues, such as [Science](#), [Ethics](#), and [Oxford University Press](#). It has been [lauded](#) by the renowned philosopher [Peter Singer](#).
- Senior Researcher [Bob Fischer](#) presented the work to the [Humane and Sustainable Food Lab](#) at Stanford University, to the [Shrimp Welfare Project](#), and others.
- Bob discussed the project on podcasts, such as [80,000 Hours](#) and [How I Learned to Love Shrimp](#).
- Other guests on each of these podcasts, including [Jeff Sebo](#) (NYU Mind, Ethics, and Policy Program) and [Amanda Hungerford](#) (Open Philanthropy), also cited the project. The latter attributed the research with changing their views on insect welfare.
- The Moral Weight Project has started to result in concrete shifts in funders' resource allocation prioritizations. For example, Animal Charity Evaluators are now integrating elements of the work into their evaluation criteria. They plan to further revise their approach in consultation with RP researchers.
- Broader audiences, such as government agencies in the United States and the Netherlands are exploring how the research might improve their decision-making.

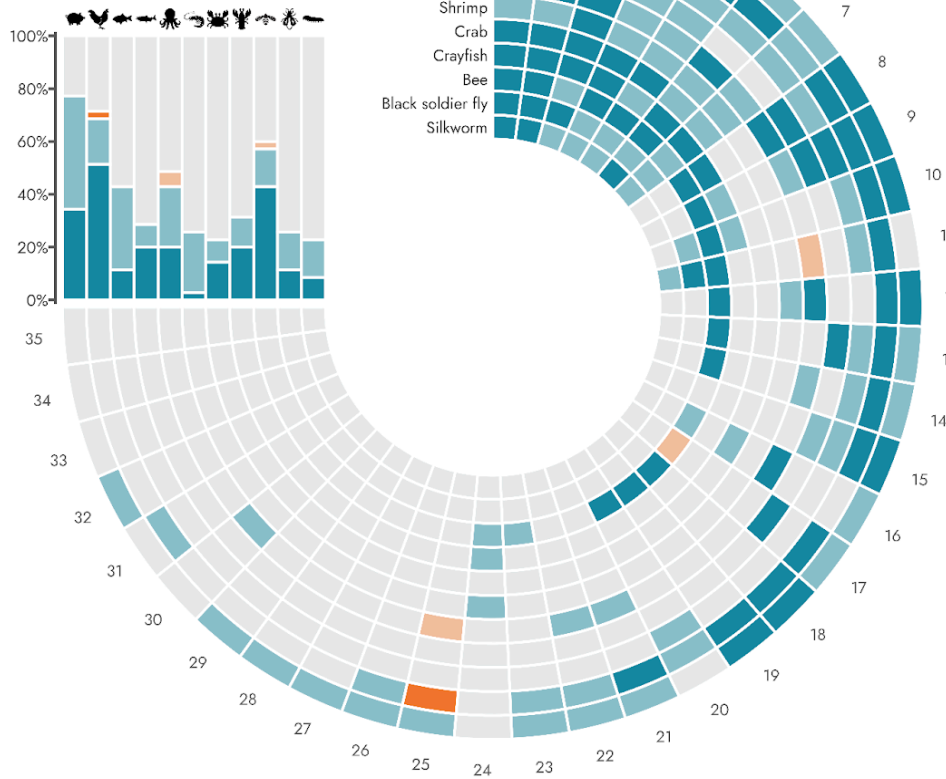
"I think the big animal-related view I've changed my mind on fairly recently is insects. I very much used to view them as these kind of unfeeling automatons, did not give them a second thought, was very happy to not worry about their wellbeing. But after seeing some of the Moral Weight Work that Rethink Priorities has put out, I now feel worried about the insects."

— Amanda Hungerford (Open Philanthropy)

The Moral Weight Project: cognitive proxies welfare range table

Cognitive proxies

Colored segments around the arc show the assessment made for each cognitive proxy in each animal. The barplot shows the cumulative percentage of proxy assessments for each animal.



Capacity present?: ■ Likely yes ■ Lean yes ■ Lean no ■ Likely no ■ Unknown

- | | | |
|---|----------------------------------|--------------------------------------|
| 1. Communication | 13. Socio-spatial cognition | 25. Mirror mark recognition |
| 2. Navigation strategies | 14. Inhibitory control | 26. Mental Time Travel |
| 3. Individual differences/"personality" | 15. Object permanence | 27. Body awareness |
| 4. Motivational trade-off | 16. Tool use | 28. Experience projection |
| 5. Reversal learning | 17. Transitive inference | 29. Perspective taking |
| 6. Multimodal integration | 18. A not B detour test | 30. Reverse reward test |
| 7. Cooperative behavior | 19. Gratification delay | 31. Symbolic representation of world |
| 8. Responses to novelty | 20. Numerical cognition | 32. Theory of mind |
| 9. Social learning | 21. Physical reasoning | 33. Memory bias |
| 10. Judgment bias | 22. Attention bias | 34. Response-slowing |
| 11. Cross-modal learning | 23. Temporal/spatial discounting | 35. Shared intentionality |
| 12. Problem-solving | 24. Uncertainty monitoring | |

Source: *Rethink Priorities*

>> [See all the welfare range tables.](#)

Building better models

The Moral Weight Project's outcomes showed that there is demand for work that makes philosophical, methodological, and empirical problems manageable, modelable, and actionable. With this goal in mind, RP established the [Worldview Investigations Team](#). To date, the team has developed three open-source tools to spur more transparent deliberations.

How can donors compare the cost-effectiveness of projects across cause areas?

The [cross-cause cost-effectiveness model](#) allows users to model their philosophical assumptions and preferences and see the concrete range of outcomes of donation choices. This project also considered alternatives to expected value maximization.

“The fund managers were highly impressed by the ambitious scope of [the cross-cause cost-effectiveness model]. Despite the effective altruism (EA) movement existing for over a decade, there were no other publicly available cross-cause models with comparable breadth and an EA-informed perspective.”

Linchuan Zhang (EA Infrastructure Fund)

How might empirical assumptions and risk attitudes impact your donations?

The [portfolio builder](#) evaluates the cost-effectiveness of different cause areas using data-driven estimates and simplified models of costs. The tool can:

1. Suggest an optimal mix of donations based on user-selected criteria.
2. Evaluate donation strategies under different decision-making approaches.

The portfolio builder aims to help users explore different ways to divide their donations that could maximize their impact.

Given that the portfolio builder and moral parliament tool were just released, it is difficult to fully assess their impact. Thus far, the team's talks on how to use the tools have been well-attended, with participants raising in-depth questions about practical applications.

How should you allocate philanthropic resources in the face of moral uncertainty?

By using the [moral parliament tool](#), users can model different worldviews as delegates, with each representing a set of ethical beliefs and values. Users can:

1. Input their confidence in various worldviews.
2. Explore methods for reaching decisions about charitable giving.

The moral parliament shows the impact of different philosophies and decision-making approaches in philanthropy.

Enabling data-driven decision making

Organizations across cause areas rely on data to maximize their impact. Maintaining specialized in-house researchers with the expertise to quickly gather and analyze survey data can be resource-intensive, especially for small nonprofits with limited budgets whose needs may arise on an ad-hoc basis.

Rethink Priorities' [Surveys and Data Analysis Team](#) fills this need.

The research team supports the broader field, offering both paid services and pro-bono work depending on their client's budget and scope of the project. Maintaining a fast turnaround (sometimes a few days) allows RP to help organizations respond more quickly to emerging opportunities or pressing issues.

Data propels impact

Knowing what specific groups of people (e.g. students, voters, etc.) think about an issue helps organizations to develop proposals that will have the greatest likelihood of success. For example, understanding opinions regarding animal welfare can help advocates to **create effective corporate campaigns or improve legislative asks**.

Understanding the nuanced opinions that various groups hold also allows organizations to create optimal messaging and communications strategies to **effectively rally the support necessary to solve the world's problems**.

Beyond researching opinions, RP's survey research also examines behaviors.

Planning for the future by understanding the present

The continued development of large language models (LLMs) will open the door to both benefits and risks. **What might be the real-world outcomes of future systems built from LLMs?** Answering this question requires understanding how people are *currently* using LLMs—a question that RP is researching with a [grant](#) from Open Philanthropy.

This work entails interviewing and surveying:

- Power users (people who frequently use LLMs for advanced purposes)
- Workers in relevant industries
- The general public

While this project is still in progress, the team expects their research to help inform future work led by the client and others looking to understand **how this emerging technology could transform our world**.