

Aligning Social Media Feed Rankings with Users' Values

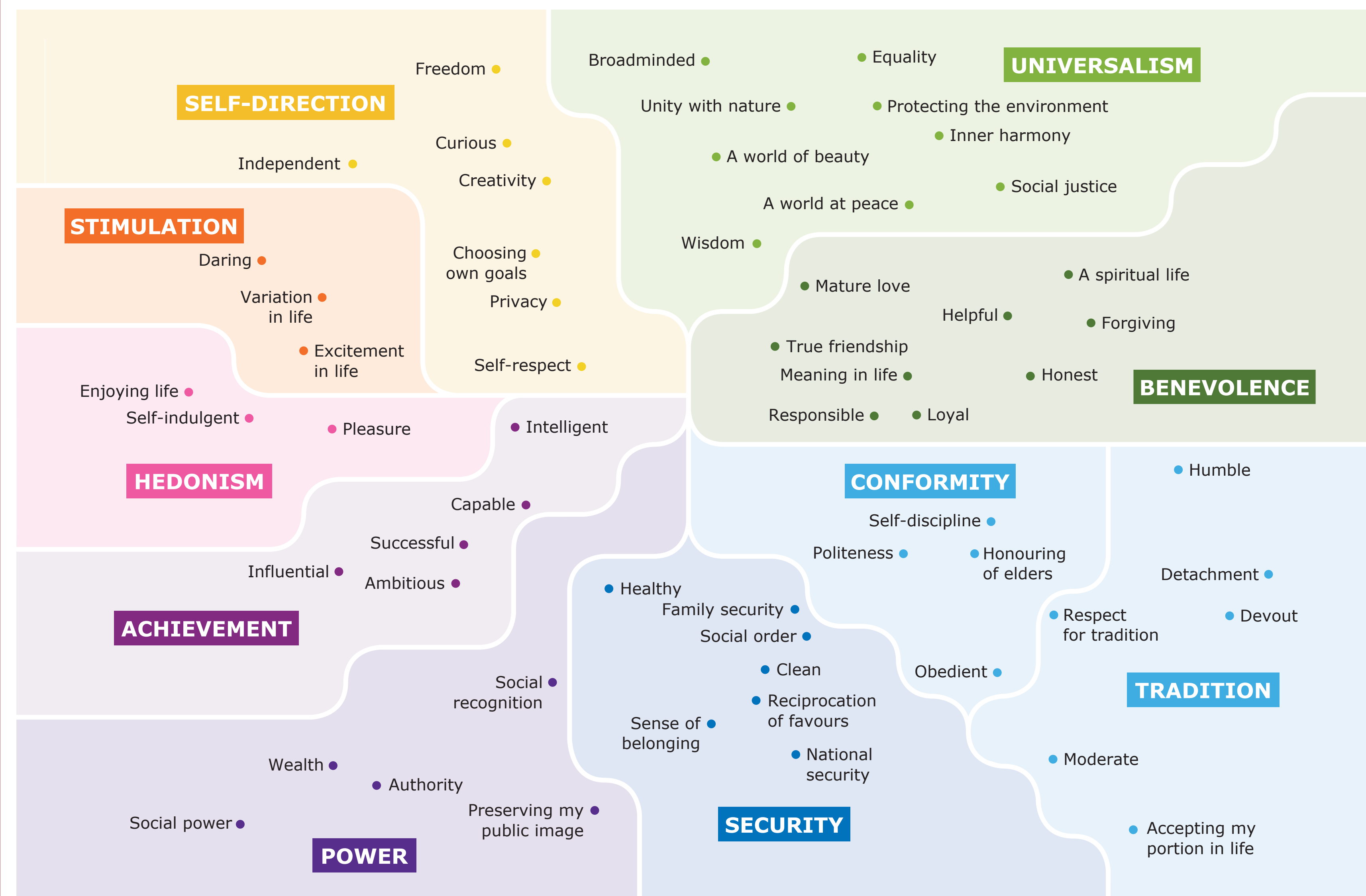
Farnaz Jahanbakhsh, Dora Zhao, Tiziano Piccardi, Ziv Epstein, Zachary Robertson, Sanmi Koyejo, Michael S. Bernstein

Social media feeds that are delivered to users embed certain values. But these values are not necessarily aligned with the values that users want to see in their feed.

Our vision: give users value-consistent feeds using tools that they can control.

How do we measure values anyway?

We borrow the Schwartz's Theory of Human Values from Social Psychology, which provides validated instruments for capturing personal and societal values.

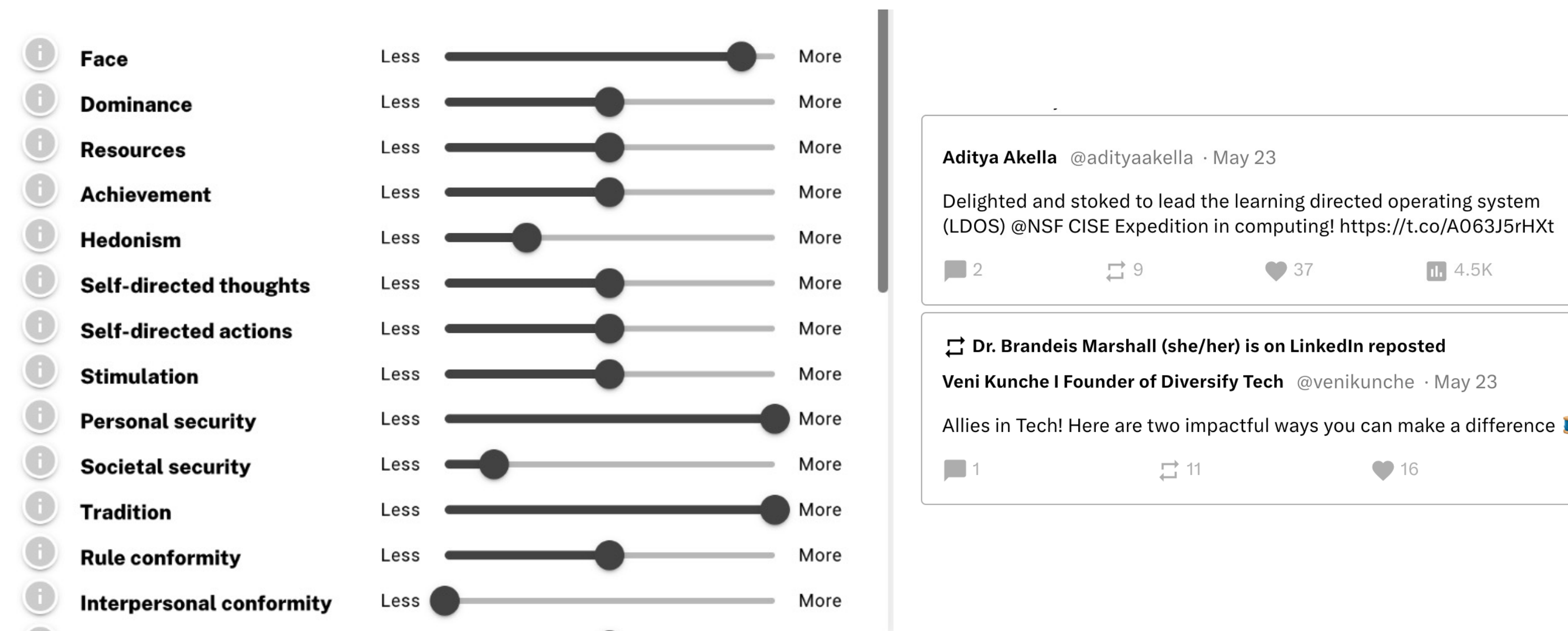


Redrawn with permission for common cause from Schwartz, S.H. (2006).

Reranking feeds based on values

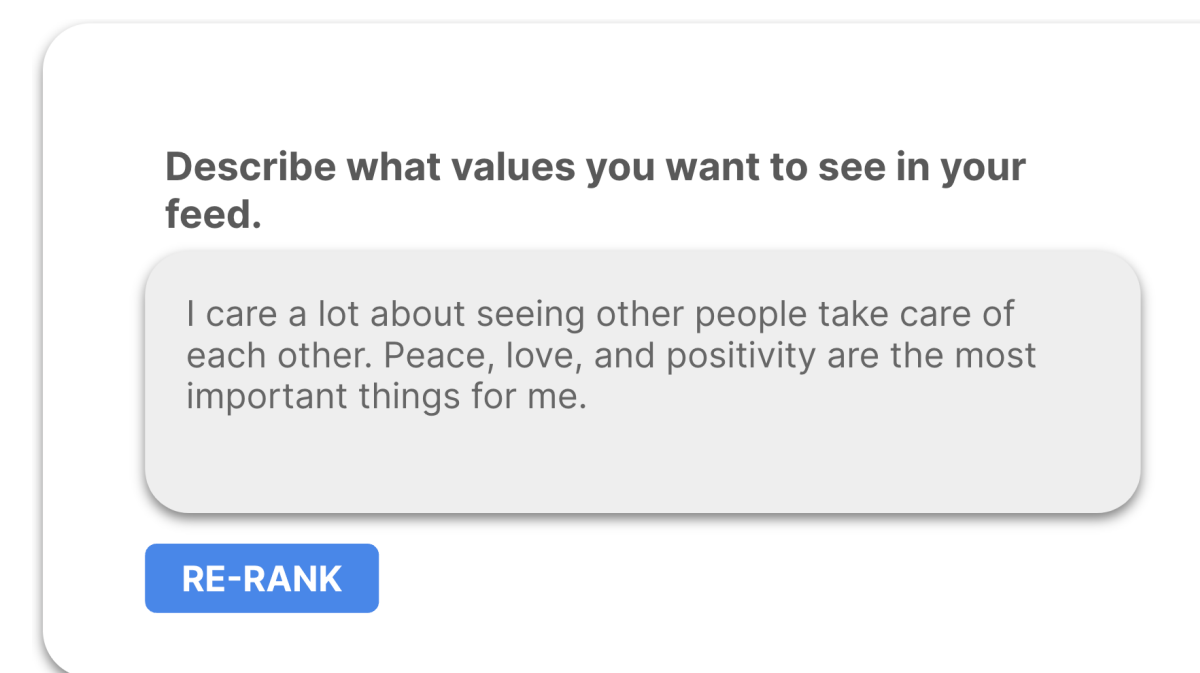
We use a classifier (*Kiesel et al. ACL'22*) to detect Schwartz's values in social media posts and offer users a suite of interfaces that enable them to specify which values they want represented in their feed.

The interfaces range from more expressive, albeit more complex:



to less expressive, but more simple:

Using natural language:

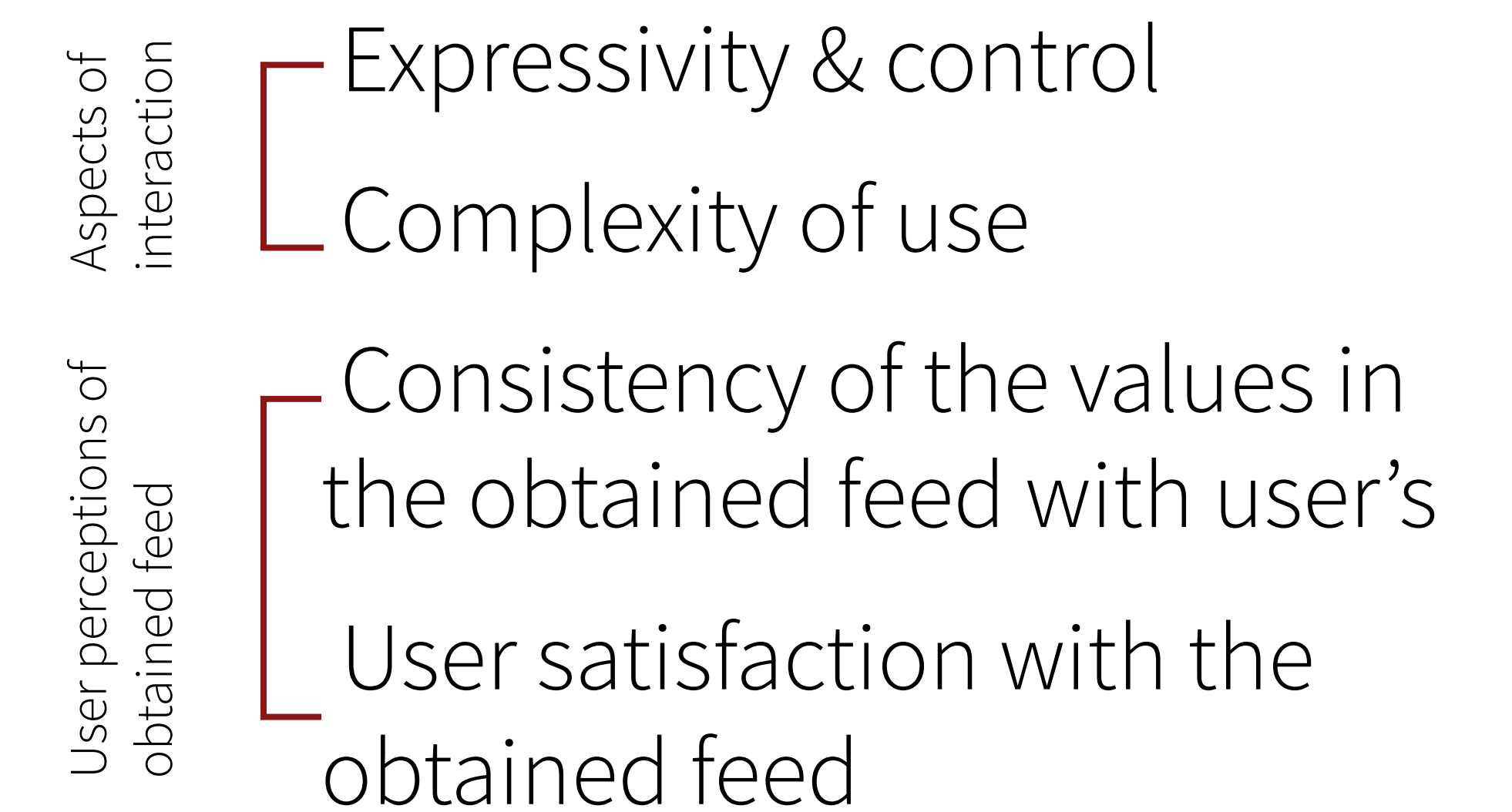


Value clusters that a user can choose from, e.g.:



Characterizing the design space

Conducting experiments to understand how these interfaces vary along:



Studying long-term effects

- **Field deploy** this tool as a browser extension so users can re-rank their social media feed in real-time based on values they deem important.
- Study how reranking feeds to prioritize different values affect various outcomes including user's personal values, emotions, and post engagement.