A Universal Moral Grammar (UMG) Ontology

Michael DeBellis
Semantics 2018
mdebellissf@gmail.com
https://tinyurl.com/UMG-Ontology-2018

What is a UMG?

- First defined by Marc Hauser in his book Moral Minds[1]
- *"I argue that our moral faculty is equipped with a universal moral grammar, a toolkit for building specific moral systems. Once we have acquired our culture's specific moral norms… we judge whether actions are permissible,... without conscious reasoning and without explicit access to the underlying principles."*
- Analogous to Universal Grammar (UG) in Computational Linguistics
- A hypothesized cognitive module that is part of the human genome

UMG Can Be Set Theoretic

- The term Grammar is used only to highlight the analogy with UG
- Indeed, even regarding UG Chomsky has stated: “...the original formulations of transformational grammar were set-theoretic, not graph-theoretic: trees are simply a pedagogical aid”[2]
- Web Ontology Language (OWL) and Semantic Web Rule Language (SWRL) are excellent tools to model a formal UMG

What is a Cognitive Module?

- First hypothesized by Chomsky (Language Faculty)[3], Marr (Vision)[4] and Fodor [5]
- Chomsky: "[domain specific mental representations] ...what you might think of metaphorically as "mental organs" on the analogy to organs of the body"[3]
- The opposite of the "blank slate" model that hypothesizes one generic mechanism for learning and reasoning
- In the last decade has been widely embraced by the Evolutionary Psychology community as a model to explain many faculties of the human mind: Morality, Theory of Mind, Living Things, Locations and Navigation

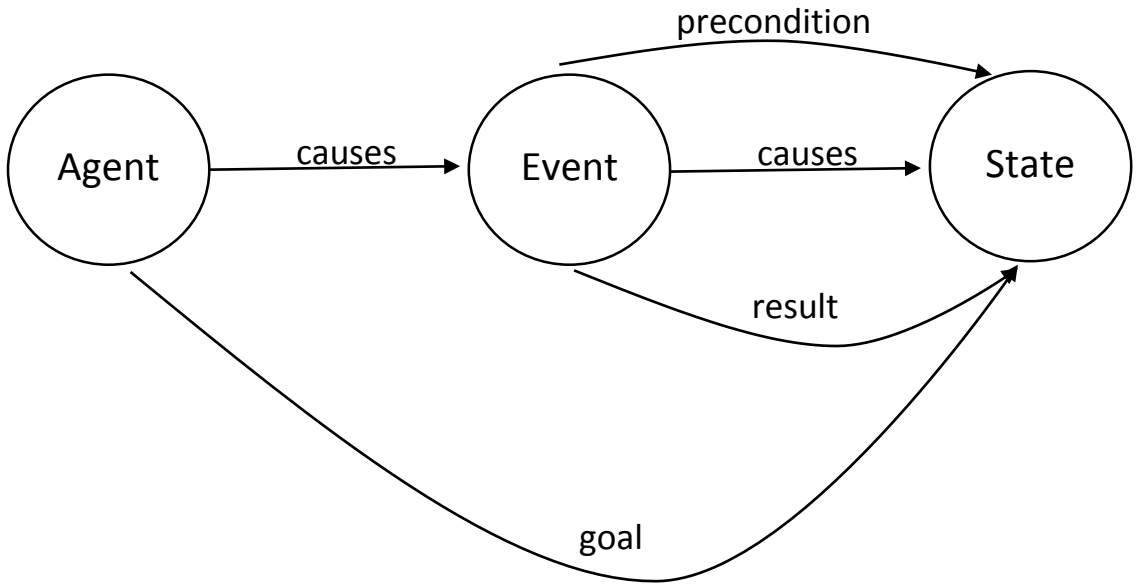
Formal Models of Cognitive Modules

- While there have been formal models of the Language Faculty and Vision there have been no formal models of these other modules except for the work of John Mikhail[6]
- Mikhail has developed a model of a UMG based on English sentences. His model is analogous to "surface structure" in computational linguistics (the syntax of a specific language) where as mine is analogous to deep structure (the hypothesized underlying model used for all natural languages)
- Many researchers describe informal "models" for these other modules
  - Theory of Mind [7]
  - Moral Faculty[7]
  - Living things[8]
- My hope is that this work is a starting point, not just to formally model the Moral Faculty but other modules as well and their interaction
- In order to model a UMG it was necessary to model various aspects of these other modules as well, especially Theory of Mind

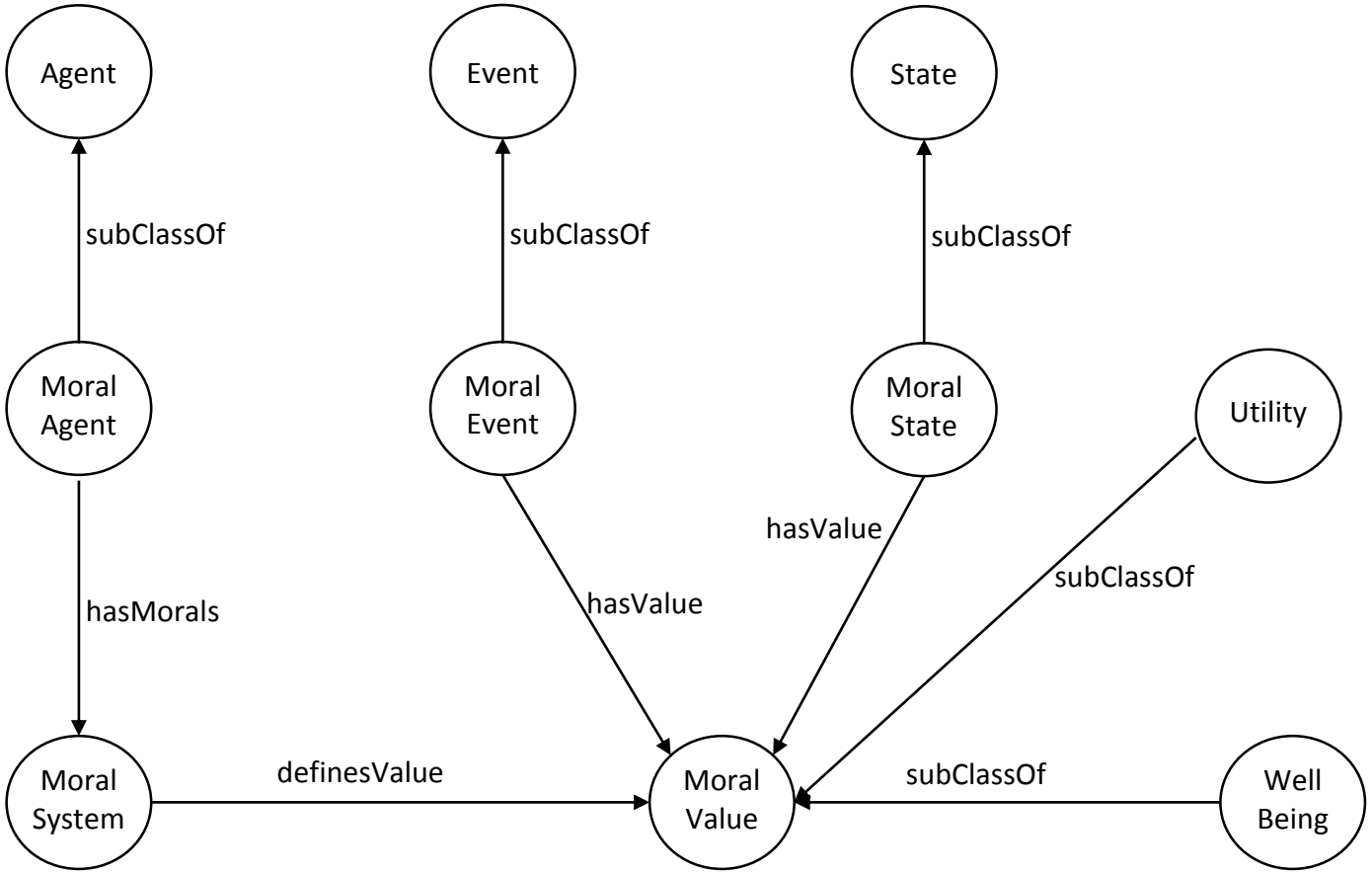
The Starting Point: Theory of Mind

- Describes Agents, Events, Causality
- Evidence for existence in pre-verbal infants[7]
- Used standard AI model for events first developed by Newell & Simon[9] which has been used for many AI systems such as SHRDLU, SOAR, and the Knowledge-Based Software Assistant (KBSA)

Fundamental Theory of Mind Model



Extending TOM to Create a Moral Model

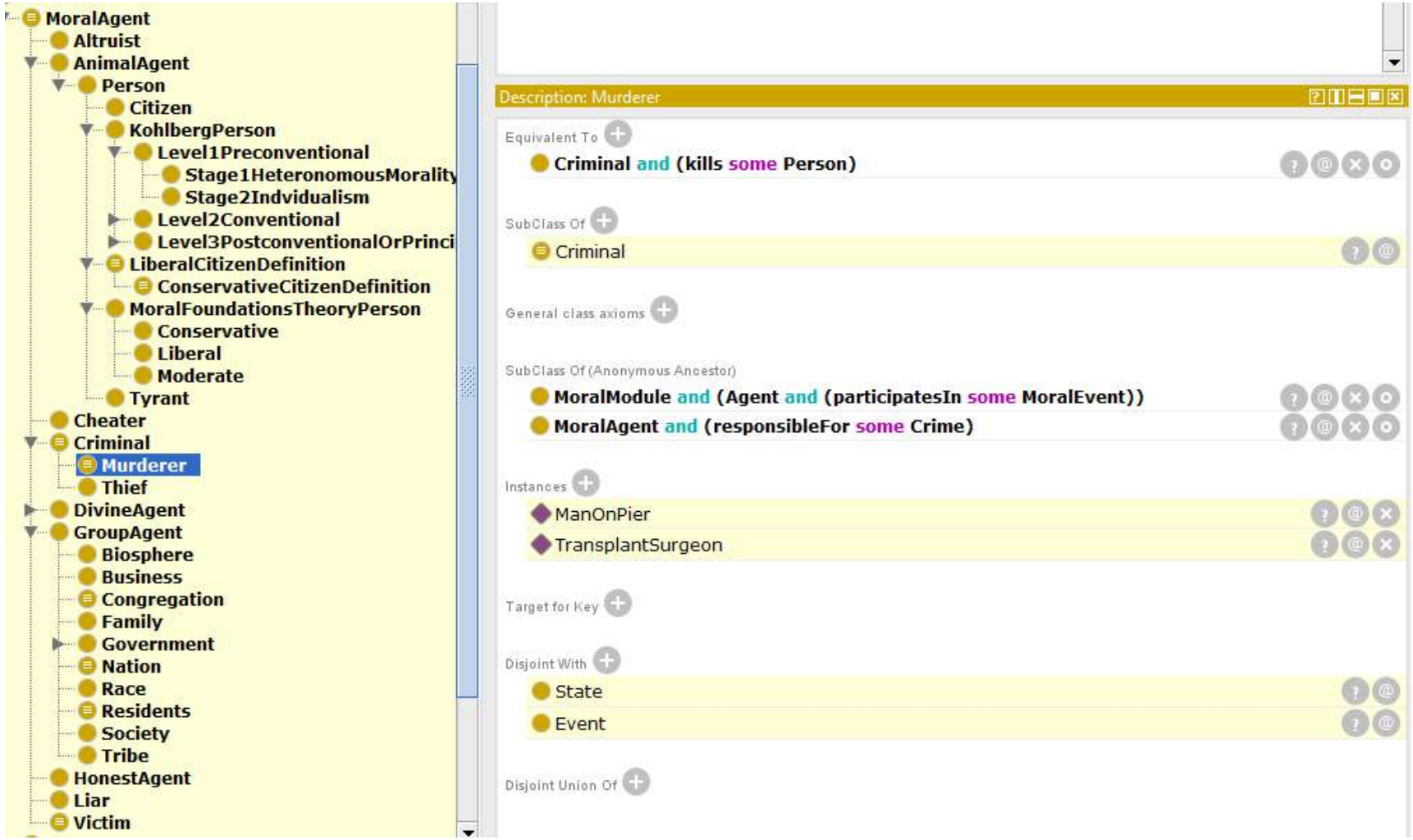


Moral Agent: An Agent that places value on Events and States

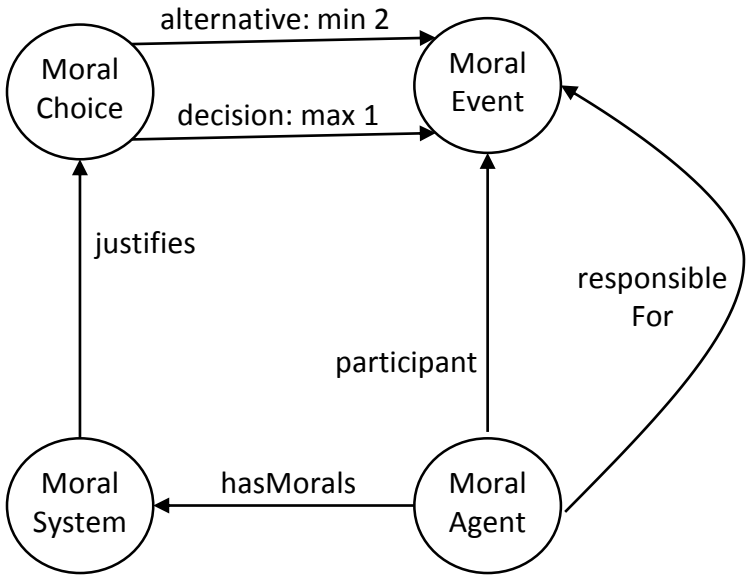
Moral Event: An Event that some Moral Agent places value on

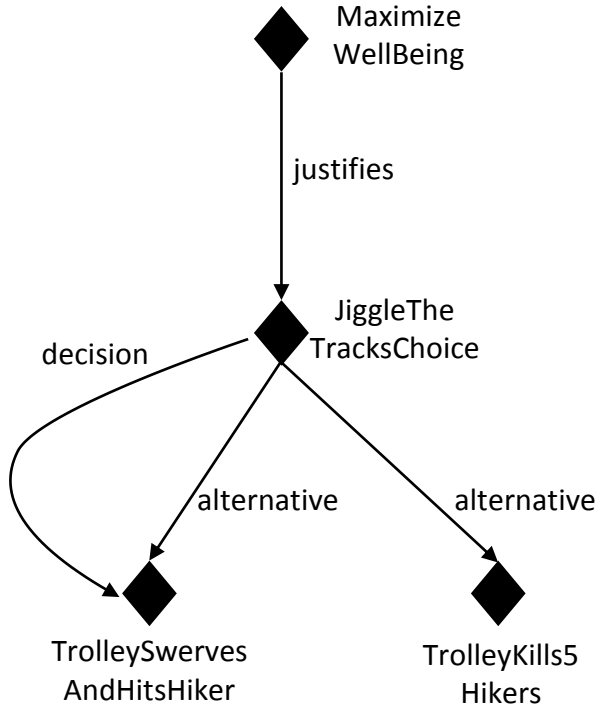
Moral State: A State that some Moral Agent places value on

Example: Moral Agent Subclasses



Important subclass of MoralEvent: MoralChoice



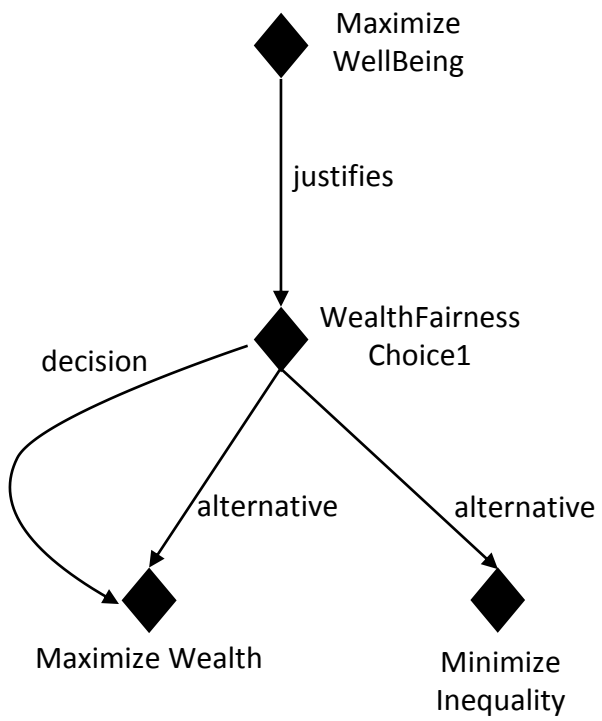


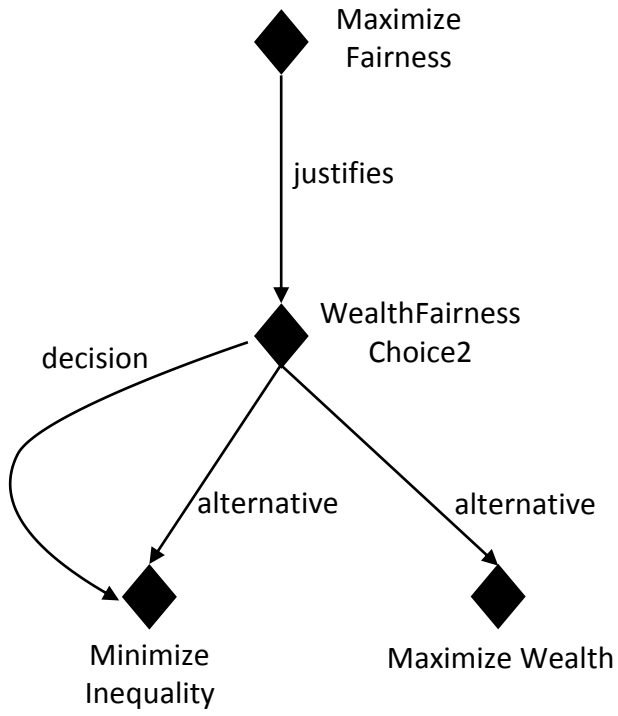
State of Current Ontology

- Implemented over 40 scenarios (Moral Choices) from the philosophical, psychological, anthropological, and biological literature
  - Fischer and Ravizza[11]
  - Marc Hauser[1]
  - Moral Foundations Theory[12]
  - Christopher Boehm[13]
- SWRL Rules that define diverse moral systems
  - Utilitarianism
  - Justice as Fairness
  - Moral Foundations Theory (MFT)
  - Golden Rule
  - Categorical Imperative
  - Fair collaboration/Free Rider identification
  - Religious dogmas
  - Kin selection

Moral Choice: Maximize Fairness or Well Being

Moral Event	Maximize Wealth	Minimize Inequality
Well Being for 95%	-1	4
Well Being for 5%	100	4
Mean Well Being	4.05	4
Standard Deviation of Well Being	22.01	0





SWRL Rules for Utilitarianism and Maximizing Fairness

- Utilitarianism: MoralChoice(?c) ^ alternative(?c, ?a1) ^ alternative(?c, ?a2) ^ result(?a1, ?r1) ^ result(?a2, ?r2) ^ meanWellBeing(?r1, ?r1mwb) ^ meanWellBeing(?r2, ?r2mwb) ^ justifiedBy(?c, MaximizeWellBeing) ^ greaterThan(?r1mwb, ?r2mwb) -> decision(?c, ?a1)
- Maximize Fairness: MoralChoice(?c) ^ alternative(?c, ?a1) ^ alternative(?c, ?a2) ^ result(?a1, ?r1) ^ result(?a2, ?r2) ^ standardDeviationOfWellBeing(?r1, ?r1sdwb) ^ standardDeviationOfWellBeing(?r2, ?r2sdwb) ^ justifiedBy(?c, MaximizeFairness) ^ lessThan(?r1sdwb, ?r2sdwb) -> decision(?c, ?a1)

Conclusion

- The UMG ontology shows that semantic technology can add rigor to the soft sciences
- The current model is a broad but shallow implementation of some of the most important examples from the philosophical, psychological, anthropological, and biological research on ethical theory
- However, the current model provides a resolution to one of the most important issues in ethical philosophy: the Is-Ought problem. My resolution is identical to the approach developed independently by Bayer and Figdor in their recent book: Atheist Heart, Humanist Mind.
- In future work I plan to use the UMG ontology to develop more sophisticated models with game theoretic analysis and/or simulations to develop a theory for the origin of human morality based on research about hunter gatherer norms. I may integrate the UMG ontology with Boehm's database of hunter gatherer tribal norms
- The UMG ontology could also be a foundation for defining norms and constraints for autonomous computer agents (i.e., e.g., Asimov's 3 laws of robotics)

Thank You!

For more info contact me at:

mdebellissf@gmail.com

And see my blog page with the ontology and an extended paper:

https://tinyurl.com/UMG-Ontology-2018

Bibliography

1. Hauser, Marc (2009). Moral Minds (P.S.) HarperCollins. Kindle Edition.
2. Chomsky, Noam (2016). Why Only Us: Language and Evolution. With Robert C. Berwick. The MIT Press; Reprint edition (January 15, 2016)
3. Chomsky, Noam (1984). Modular Approaches to the Study of the Mind. San Diego State University Press.
4. Marr, David (2010). Vision: A Computational Investigation into the Human Representation and Processing of Visual Information. MIT Press.
5. Fodor, Jerry (1983). The Modularity of Mind: An Essay on Faculty Psychology. MIT Press.
6. Mikhail, John (2007). Universal Moral Grammar: Theory, Evidence, and the Future. Trends in Cognitive Sciences. Vol. 11 No. 4..
7. Hirschfeld, Lawrence, A. (1994). Mapping the Mind: Domain specificity in cognition and culture. Edited by Lawrence A. Hirschfeld and Susan A. Gelman. Cambridge University Press. Several essays deal with modules of the type discussed in the presentation.
8. Medin, Douglas L. (1999) Folkbiology. Edited with Scott Atran. The MIT Press.
9. Newell, Allen (1958). Report on a General Problem-Solving Program. With J.C. Shaw and H.A. Simon. Proceedings of the International Conference on Information Processing. 30, December 1958. The Rand Corporation. http://bitsavers.informatik.uni-stuttgart.de/pdf/rand/ipl/P-1584_Report_On_A_General_Problem-Solving_Program_Feb59.pdf
10. Miller, George (1960). Plans and the Structure of Behavior. With Eugene Galanter and Karl H. Pribram. Holt, Rinehart, and Winston, Inc.
11. Fischer, John Martin (1952). Ethics: Problems and Principles. With Mark Ravizza. Harcourt Brace publishers.
12. Haidt, Jonathan (2012). The Righteous Mind: Why Good People Are Divided by Politics and Religion. Vintage. Kindle Edition.
13. Boehm, Christopher (2012). Moral Origins: The Evolution of Virtue, Altruism, and Shame. Basic books.
14. Carruthers, Peter (2006). The Architecture of the Mind. Oxford University Press.
15. Yang, Fan (2018). In Defense of the Commons: Young Children Negatively Evaluate and Sanction Free Riders. Psychological Science. July 16, 2018.