

The cognitive science of caregiving

Alison Gopnik (gopnik@berkeley.edu)

Department of Psychology, University of California, Berkeley, CA 94720

Max Kleiman-Weiner (maxkw@uw.edu)

Foster School of Business & Paul G. Allen School of Computer Science and Engineering,
University of Washington, WA 98105

Reut Shachnai (reut.shachnai@yale.edu) and Julia Anne Leonard (julia.leonard@yale.edu)

Department of Psychology, Yale University, New Haven, CT 06511

Lindsey Powell (ljpowell@ucsd.edu)

Department of Psychology, University of California San Diego, CA 92093

Keywords: caregiving; computational models; social cognition

Goals and Scope

Caregiving is essential to human survival and flourishing, yet it has been largely overlooked across scientific disciplines, including economics, philosophy, politics, and, importantly – cognitive science. Caregiving remains poorly understood, in part, because it does not fit neatly within traditional frameworks of human cognition and behavior (Gopnik, 2023). Take, for instance, theories of morality and cooperation (Kleiman-Weiner, Saxe, & Tenenbaum, 2017; Powell, 2022). Unlike the principle of universalism, which emphasizes impartiality, caregiving involves prioritizing the needs of specific individuals over others (Gilligan, 1993). Caregiving also directly challenges the utilitarian principle of “the greatest good for the greatest number”, since it involves actions that benefit others at significant personal cost. Furthermore, caregiving diverges from the principle of reciprocity – a cornerstone of human cooperation – because caregivers generally do not expect anything in return for their actions (Fiske, 1992).

The failure of existing theories to capture the psychological and behavioral complexities of caregiving calls for new theoretical, empirical, and computational frameworks that examine the unique cognitive mechanisms underlying the process of care.

In this symposium, we bring together leading cognitive, computational, and developmental researchers whose work aims to establish a new science of care. Alison Gopnik will start the symposium by presenting an intuitive theory of care that distinguishes it from other types of social relationships. Max Kleiman-Weiner will then present a computational framework that formalizes the goal of caregiving as promoting a learner’s future utility when the caregiver is no longer present and the learner must act autonomously. Next, Reut Shachnai and Julia Leonard will present computational and empirical evidence of the cognitive computations underlying parents’ decision to intervene (or step back) when children face challenges, with a particular focus on why parents tend to over-intervene. Finally, Lindsey Powell will discuss children’s understanding

of the complexity of care, namely, that caregivers’ seemingly antisocial behaviors (e.g., hindering, withholding of help) often have prosocial intentions. The talks will conclude with a roundtable discussion, where the presenters will draw connections across their work and identify fertile avenues for future research.

Talk 1: The intuitive theory of care

Alison Gopnik

We articulate a general account of care that distinguishes it from other social relationships. This account underpins our everyday intuitions about care and can be thought of as a distinctive intuitive theory, analogous to intuitive psychology or physics. First, care is intrinsically asymmetric, it depends on the idea that the carer has resources that the cared-for person does not. Second, care is intrinsically altruistic – it involves the carer donating resources to the cared for, regardless of return, and doing so precisely because of the cared-for’s lack of resources. There are also distinctive features of the motivations and objectives of care. Carers might be motivated to increase the objective utilities of the cared-for. But in other cases the carer may be more concerned with the subjective utilities of the cared for – what they think of as their best interests rather than what the carer might consider to be in their interest. A third form of care involves neither type of utility but rather tries to donate resources in a way that confers autonomy on the cared-for. Rather than trying to fulfill particular goals, subjective or objective – the carer works to give the cared-for enough resources to fulfill those goals themselves. We also present evidence that even young children conceive of care in ways that suggest that they have elements of this intuitive theory.

Talk 2: A computational foundation for caregiving

Max Kleiman-Weiner

I present a computational framework for understanding the principles of caregiving, addressing a critical gap between theories of altruistic helping and the unique challenges of care. While existing models of prosocial behavior represent

caregiving as optimizing the learner’s immediate utility, this approach can paradoxically undermine learner autonomy. I propose that effective caregiving instead aims to maximize the learner’s expected utility during a future period of independence after the caregiver is no longer present.

I formalize this principle using a simplified Markov decision process with two periods: a developmental phase where the caregiver and learner interact and an autonomous phase where the learner must act independently. This framework helps explain why optimal caregiving often involves creating opportunities for learning rather than directly fulfilling the learner’s immediate goals and why some exposure to manageable risks may better prepare learners for future challenges than complete protection from harm.

While optimal caregiving strategies in this framework are generally intractable, I discuss how recent advances in machine learning – particularly in curriculum design, safe exploration, and empowerment maximization – may provide approximate solutions. We test this framework empirically with behavioral studies that test our computational framework. Finally, I consider the possibilities and risks of artificially intelligent agents that might someday care for us.

Talk 3: Why parents do too much

Reut Shachnai & Julia Anne Leonard

Young children encounter challenges every day, from tying shoes to making friends. In these moments, parents must decide: should I take over or let my child figure it out? Striking the right balance is difficult—parents want children to acquire skills but also wish to prevent frustration or wasted effort if tasks are simply too hard (Shachnai, Asaba, Hu, & Leonard, 2025).

In this talk, we present new computational and empirical work uncovering how parents navigate this tension, focusing on why parents in the U.S. and other developed countries often intervene in ways that exceed children’s developmental needs (Doepke, Sorrenti, & Zilibotti, 2019). Unlike past work conceptualizing parental intervention as a stable trait shaped by individual dispositions and cultural norms, we propose that it is better understood as a dynamic, state-level decision-making process.

Our computational framework models this decision as a Partially Observable Decision-Making Process. Parents weigh two key factors: (1) their beliefs about their child’s abilities and (2) the utility (i.e., costs and rewards) of allowing independent effort. Data from over 400 U.S. parents and their preschool-aged children confirm that these factors shape parents’ real-time taking over behavior. Critically, we find that parents’ beliefs about their child’s abilities and the reward of children’s independent effort (e.g., learning) are often miscalibrated, leading to unnecessary intervention.

Using this framework, we also examine a striking pattern across datasets: parents with lower income consistently take over more. Ongoing research explores whether

income-related differences in beliefs about costs drive this pattern.

By providing a comprehensive framework of why parents “do too much”, this work serves as a first step toward building interventions that foster children’s independence, resilience, and motivation.

Talk 4: Children’s understanding of conflicts between helping and care

Lindsey Powell

Care is central to the lives of human children. They depend on other people to attend to their needs and to offer comfort when they are distressed. From an early age they also extend care toward others, in the forms of helping, sharing, and comforting. Underlying these experiences and actions, children possess a system of social cognition that allows them to recognize and evaluate actions that benefit or harm others. Current models of these cognitive processes focus on children’s ability to identify actions intended to promote others’ immediate goals or rewards.

However, caring for others does not always involve straightforward acts of support. Caregivers must often protectively hinder children who unwittingly pursue dangerous goals, and they must withhold help in order to foster children’s autonomy. This talk will describe research on the development of children’s ability to recognize the prosocial intentions behind care-based hindering or withholding of help. It will also offer hypotheses regarding the representations of one person’s concern for another’s broader welfare, not just their immediate goals, that could support this recognition of diverse forms of caregiving.

References

- Doepke, M., Sorrenti, G., & Zilibotti, F. (2019). The economics of parenting. *Annual Review of Economics*, 11(1), 55–84.
- Fiske, A. P. (1992). The four elementary forms of sociality: framework for a unified theory of social relations. *Psychological Review*, 99(4), 689–723.
- Gilligan, C. (1993). *In a different voice: Psychological theory and women’s development*. Harvard University Press.
- Gopnik, A. (2023). Caregiving in philosophy, biology & political economy. *Dædalus*, 152(1), 58–69.
- Kleiman-Weiner, M., Saxe, R., & Tenenbaum, J. B. (2017). Learning a commonsense moral theory. *Cognition*, 167, 107–123.
- Powell, L. J. (2022). Adopted utility calculus: Origins of a concept of social affiliation. *Perspectives on Psychological Science*, 17(5), 1215–1233.
- Shachnai, R., Asaba, M., Hu, L., & Leonard, J. A. (2025). Pointing out learning opportunities reduces overparenting. *Child Development*, 96(2), 679–690.