As humans, we communicate with each other readily and effortlessly, transmitting generations' worth of hard-won knowledge in a single conversation. Because both children and adults utilize these abilities so consistently, **it is natural to assume that we are quite good at teaching.** But in order to teach efficiently, teachers must *monitor* learners' knowledge states, and provide just the *right* data. While some recent work indicates that both children and adults regularly employ these abilities in simple interactions,[e.g.,1,2,3] other studies examining complex tutoring interactions have found the opposite effect: tutors often overestimate how much correct knowledge learners possess, and thus have difficulties assessing and closing knowledge gaps.[e.g.,4]

Bayesian computational models have begun providing rigorous normative accounts of optimal teaching, [e.g., 1] thus enabling comparisons between *optimal* and *actual* teacher performance. However, most prior work has not drawn on these recent advances. Additionally, prior work has not directly assessed the reasoning underlying teachers' responses, so **it is still unclear how teachers choose what to teach, and why they may make suboptimal decisions**. How do teachers: a) figure out which hypotheses their learner is currently considering (i.e., their *hypothesis space*), and then b) utilize that knowledge to decide what to teach? **I will integrate computational and behavioral methods to investigate both the cognitive abilities that enable us to teach efficiently, as well as where and why failures may occur.**

Study 1: **Can children and adults accurately infer and track changes in a learner's hypothesis space by observing the learner's information-search behaviors?** Adult and child (7-9y) participants will assume the role of "teachers" in a causal learning task, chosen to examine precisely those abilities that are so foundational to STEM learning. [e.g., 5] Because 7- to 9-year-old children can succeed on similar tasks as learners,[e.g., 6] it is possible they may succeed as teachers – and because peers often teach each other, we believe it is important to gauge their abilities.

Blocks with varied features will be presented (i.e., blocks might vary based on shape, size, and/or color), a subset of which will belong to the novel category "blicket." Blickets are so named for a novel causal property: when placed on top of a blicket-detector machine, they cause it to activate and play music. Therefore, learners can infer category boundaries by choosing one block at a time to place on the machine, and a teacher observing this process might infer the learner's hypothesis space from their actions. For example, if a learner selects three red blocks of different shapes to test on the machine in succession, it should be inferred that they are testing a color-related hypothesis (i.e., some variant on "red things are blickets"). This should be different than the inference a teacher might draw from watching a learner test three blocks of the same shape (but varied colors) on the machine – there, teachers might infer that the learner is testing a shape-related hypothesis (i.e., some variant on "things of this particular shape are blickets").

Participants will be taught which features determine blicketness, and be will shown a learner's information search. After viewing each block choice, teachers will be asked to provide the following information: a) the learner's hypothesis space – i.e., the *range* of hypotheses the learner may be considering, b) the hypotheses the learner is currently *prioritizing* in their exploration – i.e., which hypotheses the learner thinks are most likely to be true, c) which of the remaining blocks they'd choose to show the learner, to teach them what makes something a blicket, and d) how confident teachers are in their judgments. A cover story will be given to explain why learners don't see the teaching choices selected in c), and the task will end once teachers provide this information for every one of the optimal learner models' block choices.

As real learners' behavior is necessarily messy, we will present all participants with the search pattern of an *optimal Bayesian learner*, as defined by a computational model. This model will yield the best sequence of interventions (blocks) a learner should choose, and also the structure of the hypothesis space underlying each intervention. We will also model the responses of an *optimal teacher*, in order to compare teachers' responses to the model's teaching choices. Responses to a) and b) will be compared to the hypothesis space of the *optimal learner* model at each step, and will reveal teachers' accuracy in capturing learners' hypothesis spaces. Responses to c) will be compared to the responses of the *optimal teacher* model, to reveal whether participants make optimally informative teaching decisions. Lastly, part d) will reveal whether teachers' self-reported uncertainty relates to the correctness of their judgments. If teachers indeed tend to be most uncertain when they are actually incorrect, this will provide an avenue for potential interventions (i.e., encouraging teachers to seek out more information when they are uncertain). If participants succeed in Study 1, we will move forward with Study 2a. If children (or adults) have difficulties in Study 1, we will move forward with Study 2b.

Study 2a: **Can teachers still succeed in messier, "real-world" situations?** Although we know that both adults and children can conduct relatively efficient information searches,[e.g., 6,7] the real world is noisier than a simple optimal search model. To more closely approximate real-world situations, rather than computationally modeling learners' information searches, we will assign actual child and adult participants as "learners" in this task. We will record their series of block choices, and will ask them to explain each choice in order to gauge their hypothesis space. We will then replicate the methods of Study 1, but will show teachers these real-world search patterns, rather than model-generated patterns. If teachers still succeed in accurately capturing learners' hypothesis spaces from this information, this will be additional confirmation of teachers' abilities.

Study 2b: **If children and/or adults fail, might they succeed on a study with simpler methods?** We will replicate Study 1, but with a pared-down methodology. Teachers will still be shown a clean, optimal information search. After seeing each block choice, teachers will be asked to provide the following information: a) why the learner picked the block they did, b) what features the learner currently thinks determine blicketness, and c) and d) from Study 1. However, rather than generating a) and b) themselves, teachers will be presented with only two possible options in both a) and b), and will have to choose the option they think most accurately captures what the learner was thinking. Options will vary in the accuracy with which they relate to the model's actual states – but one choice will always be *better*. If teachers are able to succeed on this task by identifying the choice that better reflects the model's actual states, this will indicate that adults and/or children are able to capture their learners' hypothesis spaces. If teachers fail at this simpler, stripped-down task, we will explore ways in which failures might be mitigated.

While the methodology of these initial studies emphasizes simplicity, **future work** will explore teachers' abilities in more **ecologically valid situations**. For example, we will utilize actual schoolteachers as subjects, to see whether our results generalize to experienced educators. Because educators may be particularly good at utilizing these skills, their performance will allow us to narrow down factors and biases that affect non-educators' abilities. This work will also allow us to identify any points at which educators *do* have difficulties, and investigate ways to ameliorate problems. The results of this research will help identify areas where teachers have particular difficulty reasoning about what their learners know, and will explore whether there are ways for teachers to effectively correct their own mistakes, in order to make classroom teaching and learning more effective and successful.

1) Shafto, Goodman & Griffiths (2014). *Cog Psych*, *71*, 55-89. **2)** Rhodes, Gelman & Brickman (2010). *Dev Sci, 13*(3) 421-429. **3)** Gweon, Shafto, & Schulz (2014). *Proc Annu Conf Cog Sci Soc*. **4)** Chi, Siler & Jeong (2004). *Cog and Instr, 22*(3), 363-387. **5)** Gopnik & Sobel (2000). *Child Dev, 71*(5), 1205-1222. **6)** Ruggeri & Lombrozo (2015). *Cognition,* 143, 203-216. **7)** Gershman, Horvitz & Tenenbaum (2015). *Science, 349*(6245), 273-278.