The First Crank of the Cultural Ratchet: Learning and Transmitting Concepts through Language

Sahil Chopra (schopra8@stanford.edu)
Department of Computer Science, Stanford University

Michael Henry Tessler (tessler@mit.edu)
Department of Brain & Cognitive Sciences, MIT

Noah D. Goodman (ngoodman@stanford.edu)
Department of Psychology & Department of Computer Science, Stanford University

Abstract

Human knowledge accumulates over generations, amplifying our individual learning abilities. What is the mechanism of this accumulation? Here, we explore how language allows accurate transmission of conceptual knowledge. We introduce a novel experimental paradigm that allows direct comparison of learning from examples and learning from language. In our experiment, a *teacher* first learns a Boolean concept from examples; they then communicate this concept to a *student* in a free conversation; finally, we test both teacher and student on the same transfer items. We find that learning from language is both sufficient and efficient: Students achieve accuracy very close to their teachers, while studying for less time. We then explore the language used by teachers and find heavy reliance on generics and quantifiers. Taken together, these results suggest that cultural accumulation of conceptual knowledge arises from the ability of language to directly convey generalizations.

Keywords: concept learning; cultural ratchet; communication

Introduction

The human species is remarkable: We are able to learn by observing the world around us, forming new concepts that support prediction and manipulation (Tenenbaum, Kemp, Griffiths, & Goodman, 2011). Yet human concept learning has limits: Life is only so long and a person can only be in one place at a time. Individual learning from observations is thus unlikely to fully explain the ecological successes of our species (Henrich, 2015). If we are able to faithfully transmit our knowledge to the next generation, then limited individual learning can accumulate over generations to arrive at powerful systems of knowledge—a process termed the "cultural ratchet" (Tomasello, 1999). How does the ratchet work? What aspects of cognition support faithful transmission?

Cultural transmission has been often studied through the lens of imitation. This mechanism is particularly useful for learning procedural knowledge and rituals (Legare & Nielsen, 2015). Reproducing the behaviors of conspecifics, however, does not easily address ideas that go beyond the here-and-now: our generalizable knowledge and intuitive theories. Language, on the other hand, is a tool by which humans can convey abstract information. It allows us to transmit knowledge that would be otherwise difficult or unsafe to observe directly (e.g., which plants are poisonous; Gelman, 2009; Tessler, Goodman, & Frank, 2017).

Prior experimental work in cultural transmission has suggested that language may be a sufficiently expressive channel for conveying hard-to-discover knowledge (Beppu & Griffiths, 2009; Morgan et al., 2015). For example, Morgan et al. (2015) found that knowledge about stone flaking and tool making were best transmitted via verbal language. This work did not examine in detail, however, the kinds of natural language expressions utilized in the transmission of knowledge, nor relate it to the concepts being transmitted. In this paper we introduce a novel experimental paradigm that allows to explore how language can support the "first crank" of the cultural ratchet: how concepts learned from examples by one person are faithfully transmitted to a second via language.

Typical concept learning experiments are structured so that a single subject is presented with examples of objects that belong to (and don't belong to) a new category (Bruner, 1956; Piantadosi, Tenenbaum, & Goodman, 2016). We extend this paradigm by asking the initial learner to convey the concept to a second person. We allow them to do so freely using language. We then separately test the initial and secondary learner on the category. This allows us to explore detailed questions about whether and how language allows faithful transmission of these concepts: Is language sufficient for conveying concepts? How efficient is language compared to directly studying examples? What aspects of language are used to convey concepts?

In the remainder of the paper we introduce our experimental paradigm and then explore the resulting data with a variety of analyses. We find that language is sufficient and efficient for concept learning, and that certain linguistic forms seem to underlie this efficacy.

Methods

Participants

We recruited 224 participants from Amazon's Mechanical Turk (MTurk). This number was chosen to yield approximately 10 dyads per concept. Participants were restricted to those with U.S. IP addresses and at least a 95% work approval rating; in addition, participants who self-reported a native language other than English or failed to partake in the experiment (accepted the hit but then discussed matters entirely unrelated to the experiment) were excluded. In total, 11 pairs were excluded on this basis. The experiment took on average 15 minutes and participants were compensated $1.25 with an additional performance bonus (described below).

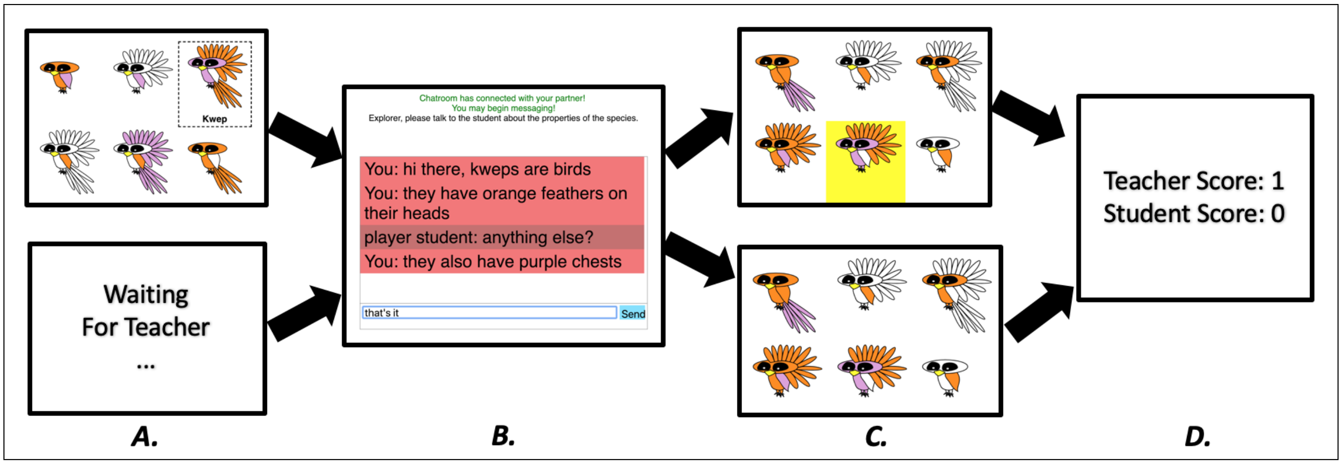

Figure 1: During the *concept learning* phase (**A**) the teacher (above) clicked through a grid of creatures, revealing the labels for creatures in the train set, while the student (below) waited. During the *concept communicating* phase (**B**), the teacher explained the concept to the student in a chatroom. During the *concept testing* phase (**C**), both participants were shown the same grid of held-out test creatures. Each selected the creatures (yellow) that they believed belonged to the concept. Finally, the participants were shown their scores for the round (**D**).

Concepts and stimuli

Participants learned concepts generated by 5 different *rules* (i.e. logical forms): Single Feature, Conjunction, Disjunction, Conjunctive Disjunction, and Disjunctive Conjunction. Rules were realized in specific *concepts* by varying Boolean properties of programatically generated images of creatures, from five different kinds: flowers, bugs, birds, fish, and trees (see Figure 1 for an example). Each kind had 5 to 7 Boolean features that we used to realize our concepts. Each of the 5 rules was realized twice in each creature kind, yielding a total of 50 concepts (listed on the axis of Figure 3). For each concept, we generated 100 specific creatures, split into 50 for training and 50 for testing. We ensured some positive examples of the concept even for very restrictive rules by first randomly selecting 6 positive instances of the concept and then adding 44 items chosen at random from all remaining items (i.e., according to the true concept base rate).

Procedure

Every pair of participants was placed in a game, where one was assigned the role of the "teacher" (initial learner) and the other was assigned the role of the "student" (secondary learner). Each game consisted of 5 rounds, each with a new concept from a new rule. Each of a game's 5 concepts used a different creature kind, and each concept was presented with a different nonce word as the species name. The ordering of concepts was randomized so that there was no standard ordering of rule types across the games.

On each round, participants went through three phases: *concept learning*, *concept communicating*, and *concept testing* (Figure 1). During the *concept learning* phase, the teacher was presented a grid of training creatures and was instructed to click on individual creatures to reveal whether or not they belonged to the species defined by the concept. Once the teacher clicked on every creature in the grid, they were presented a message advising them to review the creatures for as long as they needed. When the teacher ended the concept learning phase, they proceeded to the *concept communicating* phase, where they entered an online chatroom and were were instructed to teach the concept to the student. Participants were provided no additional instructions for the chatroom, and they were allowed to talk freely. In order to prevent a teacher from rushing through the chatroom without properly communicating with their student, only the student was given the ability proceed to the final *concept testing* phase. In the final phase both participants were (separately) given the same grid of test creatures and asked to tag the creatures that they believed belonged to the species. Neither participant had access to their chatroom messages during this phase.

Once both participants completed *concept testing* for a concept, they were provided feedback in the form of their own and their partner's score, computed as: # of hits − # of false alarms. We encouraged them to learn concepts thoroughly and communicate effectively with a monetary bonus equal to the sum of both players' scores (in cents). Participants were made aware of the task structure and bonus mechanic prior to starting the first round; they had to answer 5 comprehension questions correctly to begin to the game.

Analysis and Results

Our experiment yielded rich data for exploring whether and how concepts are learned from language, and how learning from language compares to learning from examples. We first examine performance during the *concept testing* phase for both the student and teacher participants. We then explore the time spent learning from each type of evidence. Finally, we explore the actual language used to teach concepts in the *concept communicating* phase.

Concept learning performance

Participants assigned to be the *teacher* take part in a standard Boolean concept learning paradigm, and results are in accord with expectations. The five rule types we used in our experiment cover a range of complexity in terms of description length (Feldman, 2000), which manifests in variable performance in test accuracy across types (Figure 2).

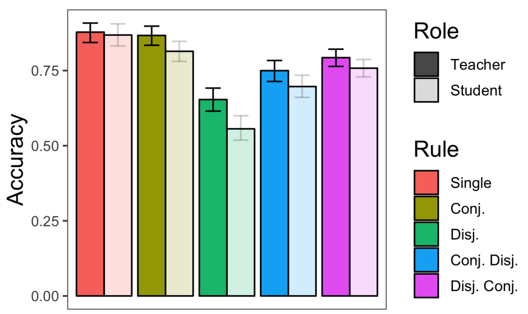

Figure 2: Average accuracy of teachers and students during the *concept communicating* phase of the experiment. Error bars denote bootstrapped 95% confidence intervals.

Students have access to the concept only through the language conveyed by their partner. The average student accuracy during the *concept testing* phase should thus be no greater than the average teacher accuracy, which appears to be true for our 5 rule types in Figure 2. To assess the potential accuracy differences between learning from examples vs. from language, we built a Bayesian mixed-effects model predicting whether or not a participant responded accurately during the *concept testing* phase as a function of the rule, the participant's role (teacher vs. student), and their interaction. We included random intercepts and effect of rule for participants and random intercepts and effect of role for each of the 50 concepts. All regression models were created in Stan (http://mc-stan.org/) accessed with the brms package (Brkner, 2017). We find a main effect of role such that students were less accurate than teachers (posterior mean and 95% credible interval: $\beta = -0.41(-0.69, -0.12)$). However, this effect is very small in absolute terms—the average difference in accuracy for students vs. teachers is just 5.3% (95% credible interval: 2.7%, 8.2%). Thus language appears to be sufficient to convey concepts; students are able to learn concepts from language, yielding performance very close to their teachers, who had access to the actual training examples.

Performance on individual concepts (rules reified in particular stimuli) reveal substantial variability in learning. Figure 3 shows the average performance of teachers and students for each of the 50 concepts along with the concept-specific chance accuracy[1]. Teachers perform above chance in all concepts, but there is significant variation in performance for concepts within a given rule. Such variation is expected given the known importance of feature salience and other stimulus properties on concept learning (Nosofsky, 1986). Notably, the gap between teacher and student performance also varies.[2] This variability cannot be attributed to stimulus features, which are shared between teacher and student, but rather reflect the language available for conveying different features. Inspection reveals that concepts with a large gap in teacher-student performance have a small number of teachers who used language in idiosyncratic ways. For example, one teacher described creatures belonging to the concept "bugs: no wings" as "like a worm ... [with a] straighten[ed] body". Another teacher described "flowers: purple petals OR thorns" as "no color ... a flower with sharp edge branches and some tails". In both cases the teacher fails to use a simple word for the relevant feature ("wings", "thorns") unlike most other participants. These cases may arise from particularly confused teachers, particularly difficult to describe features, or an interaction. We return to this question below.

Often, how well a person learns depends on the particular person they learned from. We find a strong linear relationship between average student accuracy and (corresponding) teacher accuracy across the 50 concepts ($r = .88$, $p < .001$; Figure 4). We further find that this correlation remains strong at the individual level ($r = .60$, $p < .001$; Figure 4).

While this suggests that students make mistakes when their teacher does, we may further ask whether they make the *same* mistakes. Since teachers and students are presented the same held-out test examples in the same order during the *concept testing* phase of the experiment, we can measure the similarity between teacher and student responses at the level of individual stimuli using Hamming distance (the total number of times the student and teacher responded differently). The average distance between teacher and student in our data set is 11.1 differences (out of 50 possible). To calibrate this number we computed a baseline by randomly permuting teacher-student pairings, which yields average distance 19.9 (95% CI $[19.84, 19.96]$). A second, tighter, baseline considers permutation of student-teacher pairs only within each concept (matching evidence seen by teachers). This yields average distance 13.53 (95% CI $[13.18, 13.91]$). Thus we can conclude that students' pattern of responses is more similar to their own teachers' responses than to other teachers in the same concept (and in the whole data set). Language seems to be sufficient to convey the concept as understood by the teacher, even when the teacher has learned the wrong thing.

We do not have a direct measure of teacher confidence that

[1]Chance is defined here as the accuracy achieved by guessing at random but with the base rate of positive examples shown for that concept. This is a stronger comparison than random guessing.

[2]Generally teachers do better than students. Ten concepts show the opposite trend, to varying extents. Three of these differences are driven by a few outliers where the teacher attained low accuracy in the final phase even though they properly communicated the concept. Seven of these concepts have students that are negligibly more accurate than teachers, i.e. correctly identify 1-2 more stimuli, of the 50 presented.

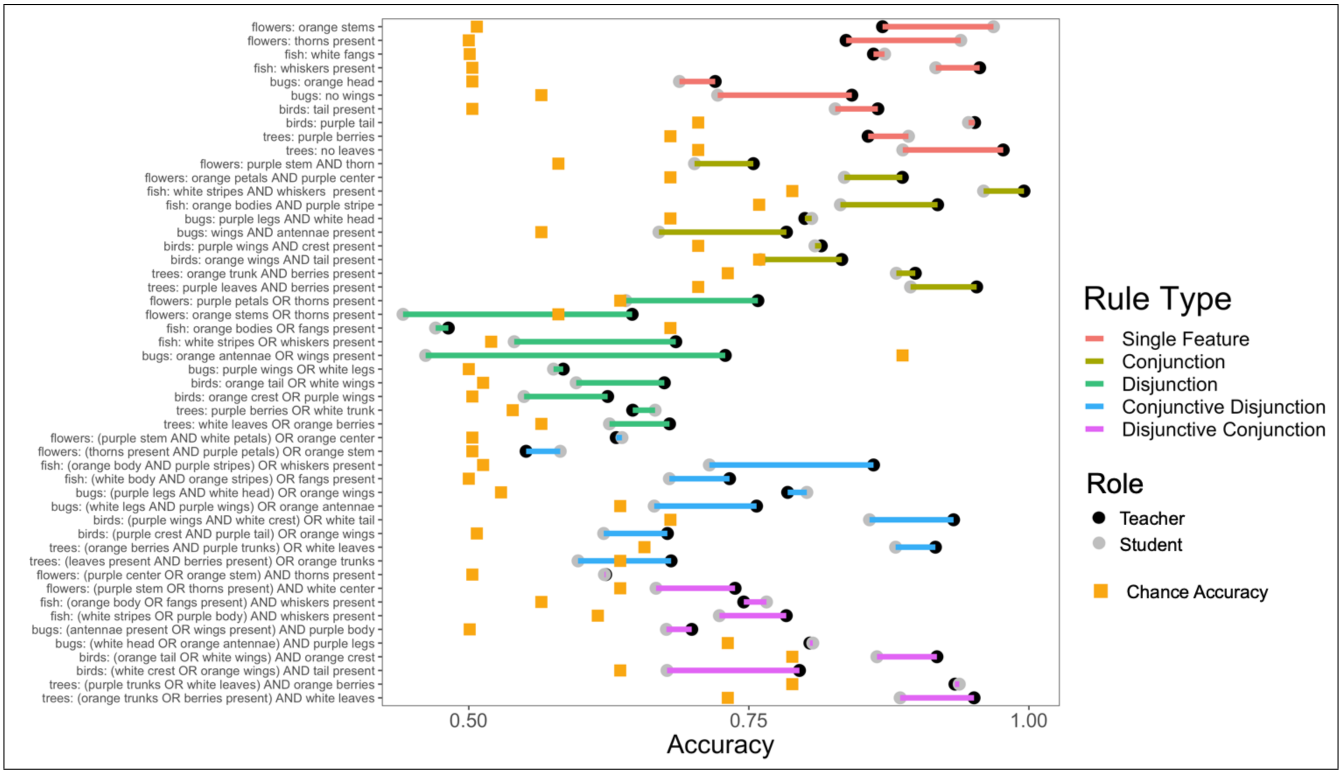

Figure 3: Accuracy on each concept. Black dots denote the average teacher accuracy on the test set; gray dots denote the average student accuracy. Gold squares denote chance accuracy.

we could use to explore the impact of confidence on the efficacy of language for transmission. Instead, we analyze an indirect measure of teacher confidence: the mean teacher accuracy within a concept. Figure 5 shows the relationship between teacher accuracy and distance from teacher to student responses. We find a strong relationship: language seems to yield stronger alignment between students and teachers when the teachers are (expected to be) confident in what they have learned ($r = -0.75, p < 001$).

Study time for observation vs. language

As we saw above, language appears to be relatively *sufficient* for conveying concepts, how *efficient* is language compared to directly learning from observed examples? We could consider efficiency in terms of amount of evidence required to learn or amount of effort required. In our experiment the amount of evidence was fixed in the *concept learning* phase, but the study time was controlled by participants. We thus consider study time as a proxy for learning effort. Since the amount of time spent in the *concept communicating* phase was similarly controlled by participants, we use time as a proxy also for effort required to convey a concept with language. Using time to measure learning effort makes it possible to directly compare effort required to learn from observations and from language.

For each concept, we recorded the amount of the time that teachers spent in the *concept learning* phase. During the *concept communicating* phase we recorded the time that elapsed between the moment a participant began typing a message into the chatbox and the moment they sent the message to their partner. Since some messages may have been unrelated to learning (e.g. pleasantries or commentary), we coded every message in the data set as “Informative”, “Follow-Up”, “Confirmation”, “Miscellaneous”, or “Social”. Informative messages were those related to the concept that were sent by teachers without prompting from the student. Messages in the ensuing dialog that were relevant to the concept were labeled as Follow-Up. Social pleasantries (“hi”, “hello”, etc.) were labeled as Social, and messages that were unrelated to the current concept (e.g. commentary about performance on previous rounds) were labeled as “Miscellaneous”. Overall, there were 1012 Informative, 1751 Follow-Up, 160 Social, and 300 Miscellaneous messages in the data set. For our time analysis, we only considered the concept-related messages: the Informative and Follow-Up messages that constituted the majority of participants’ conversations.

To compare the study time of learning from examples vs. from language (Figure 6), we built a Bayesian mixed-effects model with fixed effects of rule, participant role, and their interaction; in addition, we included random intercepts and effects of rule for each participant and random intercepts and effects of participant role for each concept.[3] We observe

[3]The data was modeled as being generated from a lognormal dis-

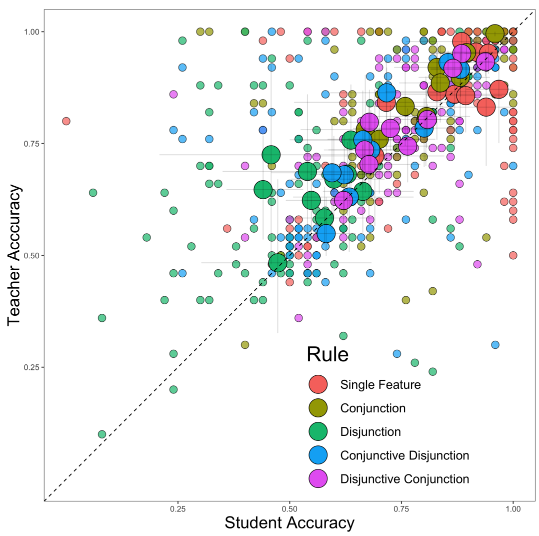

Figure 4: Accuracy of student-teacher pairs in *concept testing* phase. Small dots indicate individual teacher-student pairs, while larger dots indicate mean accuracy of teachers and students for a concept. Lines indicate bootstrapped 95% confidence intervals of teacher and student accuracy for a concept.

that the simplest rule (Single Feature) took the teacher substantially less time to study than average (comparison to the grand mean: $\beta = -0.25(-0.37, -0.12)$) and the most difficult rule took substantially more time to study than average ($\beta = 0.20(0.08, 0.31)$). Crucially, study time was systematically shorter when learning from language than from examples ($\beta = -0.64(-0.82, -0.48)$), which translated into an average 57 seconds (42, 75) less time for learning from language. There were no interactions between role and rule that were plausibly different from 0.

This suggests that learning from language may be *more*

tribution.

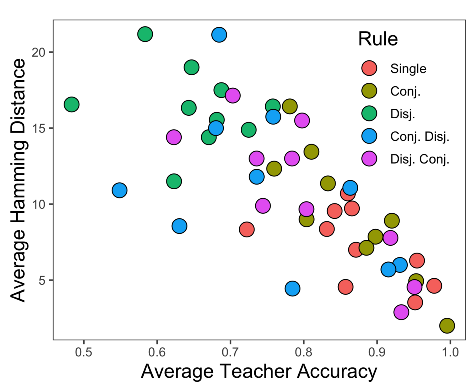

Figure 5: Average accuracy of the teacher versus the average hamming distance between student and teacher responses during *concept testing* phases of all 50 concepts.

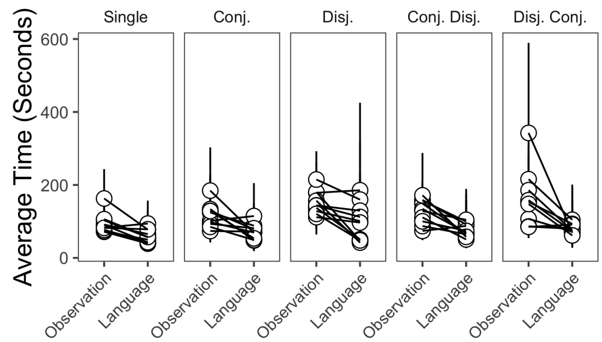

Figure 6: Time spent by teachers learning concepts from observation and time spent by teacher-student pairs communicating about concepts. Circles denote average time for a concept, error bars are bootstrapped 95% confidence intervals. Lines pair the same concept.

efficient than learning from observing examples. This conclusion warrants further study however, as our measures of study time likely depend on specific paradigm choices. For instance, teachers were forced to click on all 50 creatures during the *concept learning* phases of the experiment—it may be that not all of this time was needed for belief updating (as opposed to rote clicking of the stimuli).

Language used for knowledge transmission

We have seen that language is a sufficient and (probably) efficient means for transmitting concepts in our experiment. Now we turn to the question of what specific aspects of language were used by teachers to convey concepts. We first coded each of the messages in the game as Informative, Follow-Up, Social, or Miscellaneous, as described above. A vast majority of the messages (2763 of the 3223) were concept-relevant, i.e. Informative or Follow-up.

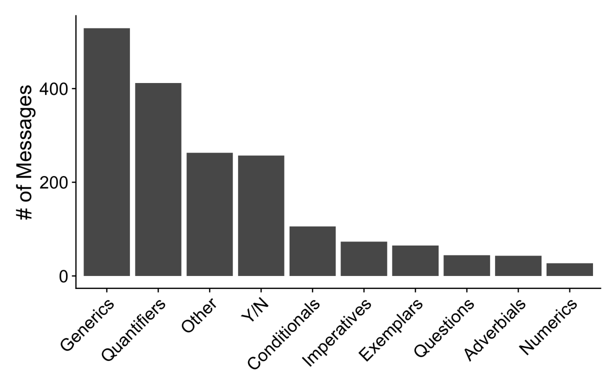

Figure 7: Distribution of concept-relevant messages.

When properties are predicated on categories, the resulting linguistic expression is typically a quantified sentence (e.g.,“All wugs have orange heads”; “Most feps have purple

Table 1: Utterance Categories & Examples

Category	Example(s)
Generics	"morseth[s] come in a variety of colors" "they have saber teeth"
Quantifiers	"some will have wings" "all morseth[s] have long whiskers"
Conditionals	"if the left wing is orange click it "
Exemplars	"12 white feathers ... no tail ..." "12 white feathers ... 5 white tail..."
Imperatives	"so click on the teeth!" "focus on orange fish..."
Adverbials	"usually their wing colors match their..." "stems are usually colored as well"
Numerics	"2/3 of them are the ones that qualify" "75% of what I clicked on was a zorb"
Yes/No	"nope", "k", "yes", "okay"
Other	"this sounds difficult" "okay idk what else to say"

wings") or a generic sentence which lacks explicit quantification (e.g., "Morseths have saber teeth" Carlson & Pelletier, 1995). Rather than talking about categories explicitly, participants could convey the actual examples they saw using numerical language (e.g., "4 of them ...") or describing individual exemplars (e.g., "white-tail with feathers, white-tail with no feathers, ...").

The first author first identified Generics consistent with other coding schemes used for generic sentences (Gelman, Goetz, Sarnecka, & Flukes, 2008), then identified Quantifiers, Numerics, and Exemplars, before grouping the remaining messages by the following linguistic constructs: Conditionals, Imperatives, Adverbials, and Yes/No statements. Remaining messages were grouped as "Other". See Table 1 for examples of messages across these categories.

Figure 7 shows label counts for concept-relevant messages in our data set. The majority of these messages use generics or quantifiers to convey information about the category, with generics being the most common. Examining this distribution within rules, we find that this pattern holds for all except disjunction, where quantifiers are more prevalent than generics. Additionally, we find that the number of generics (%G) and quantifiers (%Q) amongst concept-relevant messages does not vary appreciably across the rules: single features: (35%G, 21%Q); conjunction (34%G, 22%Q); disjunction (21%G, 26%Q), conjunctive disjunction (24%G, 21%Q); disjunctive conjunction (31%G, 22%Q).

The remainder of the responses are mostly made of other commentary about the concepts and Yes/No responses. It is important to note that teachers could have directly instructed the students what to choose (with Imperatives) or described their specific experience (e.g. "there were three morseths with blue wings"); they chose instead to use linguistic constructs that convey generalizations across categories.

Discussion

In this paper we introduce the first experimental paradigm that permits apples-to-apples comparison of learning concepts from examples and from language. We found that language is *sufficient* for faithful concept transmission, in the sense that the student who learns from language is nearly as accurate as the teacher who learned from examples (and as inaccurate, making similar mistakes). We have also seen preliminary evidence that language is *efficient* for concept transmission: that it may take less time to learn a concept from helpful language than to learn it from observations.

Most work on cultural transmission either investigates well-controlled experimental paradigms but with heavily restricted modes of transmission (e.g., sharing direct observations; Efferson et al., 2007; Kalish, Griffiths, & Lewandowsky, 2007; Griffiths, Lewandowsky, & Kalish, 2013; Kirby, Cornish, & Smith, 2008; Smith, Kalish, Griffiths, & Lewandowsky, 2008; Martin et al., 2014) or use open-ended modes of transmission (e.g., creating an instructional video) but on complex tasks where a ground-truth is difficult to establish (Muthukrishna, Shulman, Vasilescu, & Henrich, 2014; Caldwell & Millen, 2008; Morgan et al., 2015). In this paper, we chart a middle course: investigating a well-studied phenomenon (Boolean concept learning) with a relatively open-ended mode of transmission (free language production).

This allows us to perform parallel analyses of *what is being learned* and *how that knowledge is conveyed*. Recent advances in natural language processing have demonstrated potential in training and parameterizing classifiers according to language (Andreas, Klein, & Levine, 2017; Srivastava, Labutov, & Mitchell, 2018). Meanwhile, there has been a growing body of research aimed at understanding effective teaching and learning within Cognitive Science (Chi, Roy, & Hausmann, 2008; Chi, Siler, Jeong, Yamauchi, & Hausmann, 2001; Kapur, 2014). We believe that bringing the pedgagogical perspective to machine learning will be instrumental to improving models that learn from language. Importantly, our novel experimental method allows for scalable data collection of language-based instruction and provides a clear classification task, i.e. training models to learn from discourse and demonstrate understanding by predicting student responses.

In our experiment, we found substantial evidence that quantifiers and, especially, generics are used by teachers to convey their knowledge about concepts. In one sense this is unsurprising, as these linguistic constructs are *about* category generalization. Yet our results provide the first direct evidence for the connection between these aspects of language and cultural transmission of knowledge. This in turn provides initial support for a strong hypothesis about the mechanisms of knowledge accumulation: *The cultural ratchet arises specifically out of the ability of language to convey generalizations through generics and quantifiers*.

Acknowledgments

The research is (partially) based upon work supported by the Defense Advanced Research Projects Agency (DARPA), via the Air Force Research Laboratory (AFRL).

References

Andreas, J., Klein, D., & Levine, S. (2017). Learning with latent language. *arXiv preprint arXiv:1711.00482*.

Beppu, A., & Griffiths, T. L. (2009). Iterated learning and the cultural ratchet. In N. A. Taatgen & H. van Rijn (Eds.), *Proceedings of the 31st annual conference of the cognitive science society* (pp. 2089–2094).

Bruner, J. S. (1956). *A study of thinking*. New York, Wiley.

Brkner, P.-C. (2017). brms: An r package for bayesian multilevel models using stan. *Journal of Statistical Software*.

Caldwell, C. A., & Millen, A. E. (2008). Studying cumulative cultural evolution in the laboratory. *Philosophical Transactions of the Royal Society B: Biological Sciences*.

Carlson, G. N., & Pelletier, F. J. (1995). *The generic book*. University of Chicago Press.

Chi, M. H., Roy, M., & Hausmann, R. G. (2008, 03). Observing tutorial dialogues collaboratively: Insights about human tutoring effectiveness from vicarious learning. *Cognitive science*, *32*, 301-41.

Chi, M. H., Siler, S. A., Jeong, H., Yamauchi, T., & Hausmann, R. G. (2001). Learning from human tutoring. *Cognitive Science*, *25*(4), 471-533.

Efferson, C., Richerson, P. J., McElreath, R., Lubell, M., Edsten, E., Waring, T. M., … Baum, W. (2007). Learning, productivity, and noise: an experimental study of cultural transmission on the bolivian altiplano. *Evolution and Human Behavior*, *28*(1), 11–17.

Feldman, J. (2000). Minimization of boolean complexity in human concept learning. *Nature*, *407*(6804), 630–633.

Gelman, S. A. (2009). Learning from others: Childrens construction of concepts. *Annual Review of Psychology*.

Gelman, S. A., Goetz, P. J., Sarnecka, B. W., & Flukes, J. (2008). Generic language in parent-child conversations. *Language Learning and Development*, *4*(1), 131.

Griffiths, T. L., Lewandowsky, S., & Kalish, M. L. (2013). The effects of cultural transmission are modulated by the amount of information transmitted. *Cognitive Science*, *37*(5), 953-967.

Henrich, J. (2015). Culture and social behavior. *Current Opinion in Behavioral Sciences*, *3*(1), 84–89.

Kalish, M. L., Griffiths, T. L., & Lewandowsky, S. (2007, Apr 01). Iterated learning: Intergenerational knowledge transmission reveals inductive biases. *Psychonomic Bulletin & Review*, *14*(2), 288–294.

Kapur, M. (2014, 03). Productive failure in learning math. *Cognitive science*, *38*.

Kirby, S., Cornish, H., & Smith, K. (2008). Cumulative cultural evolution in the laboratory: An experimental approach to the origins of structure in human language. *Proceedings of the National Academy of Sciences*, *105*(31), 10681–10686.

Legare, C. H., & Nielsen, M. (2015). Imitation and innovation: The dual engines of cultural learning. *Trends in Cognitive Sciences*, *19*(11), 688–699.

Martin, D., Hutchison, J., Slessor, G., Urquhart, J., Cunningham, S. J., & Smith, K. (2014). The spontaneous formation of stereotypes via cumulative cultural evolution. *Psychological Science*, *25*(9), 1777–1786.

Morgan, T. J. H., Uomini, N. T., Rendell, L. E., Chouinard-Thuly, L., Street, S. E., Lewis, H. M., … et al. (2015, Jan). *Experimental evidence for the co-evolution of hominin tool-making teaching and language.* Nature Publishing Group. Retrieved from `https://www.nature.com/articles/ncomms7029`

Muthukrishna, M., Shulman, B. W., Vasilescu, V., & Henrich, J. (2014). Sociality influences cultural complexity. *Proceedings of the Royal Society B: Biological Sciences*.

Nosofsky, R. M. (1986). Attention, similarity, and the identification-categorization relationship. *Journal of Experimental Psychology: General*, *115*(1), 39-57.

Piantadosi, S. T., Tenenbaum, J. B., & Goodman, N. D. (2016). The logical primitives of thought: Empirical foundations for compositional cognitive models. *Psychological Review*, *123*(4), 392424.

Smith, K., Kalish, M. L., Griffiths, T. L., & Lewandowsky, S. (2008). Introduction. cultural transmission and the evolution of human behaviour. *Philosophical Transactions of the Royal Society B: Biological Sciences*, *363*(1509), 3469-3476.

Srivastava, S., Labutov, I., & Mitchell, T. (2018). Zero-shot learning of classifiers from natural language quantification. In *Proceedings of the 56th annual meeting of the association for computational linguistics (volume 1: Long papers)* (pp. 306–316).

Tenenbaum, J. B., Kemp, C., Griffiths, T. L., & Goodman, N. D. (2011). How to grow a mind: Statistics, structure, and abstraction. *Science*, *331*(6022), 1279-1285.

Tessler, M. H., Goodman, N. D., & Frank, M. C. (2017). Avoiding frostbite: It helps to learn from others. *Behavioral and Brain Sciences*, *40*.

Tomasello, M. (1999). *The cultural origins of human cognition*. Harvard University Press.