UC Merced
Proceedings of the Annual Meeting of the Cognitive Science Society

Title
Teaching With Examples to Communicate Conceptual Uncertainty

Permalink
https://escholarship.org/uc/item/1s4013qm

Journal
Proceedings of the Annual Meeting of the Cognitive Science Society, 48(0)

Authors
Levy, Ariel
Tenenbaum, Joshua B.
Cushman, Fiery

Publication Date
2026

Copyright Information
This work is made available under the terms of a Creative Commons Attribution License, available at https://creativecommons.org/licenses/by/4.0/

Peer reviewed

eScholarship.org Powered by the California Digital Library
University of California

Teaching With Examples to Communicate Conceptual Uncertainty

Ariel Levy[1], **Joshua B. Tenenbaum**[2] **& Fiery Cushman**[1]
ariellevy@g.harvard.edu, jbt@mit.edu, cushman@fas.harvard.edu

[1] Department of Psychology, Harvard University, Cambridge, MA, USA
[2] Department of Brain and Cognitive Sciences, MIT, Cambridge, MA, USA

Abstract

When communicating our conceptual knowledge to others, we can either provide *examples*, or provide the defining *rule* of the concept directly. Here, we propose a normative computational theory of when people should use each mode of transmission. Our model is based on the idea that teachers who are uncertain about a target concept prefer to induce a similar structure of uncertainty in the learner's mind. Simulations show that example-based teaching more effectively communicates conceptual uncertainty, whereas rule-based teaching is more efficient when teachers are confident. A pre-registered study using a "number game" paradigm (n=100) supports this account: participants assigned the role of teachers made choices that align with the model predictions when deciding whether to communicate examples or rules. Subsequent participants (n=100) that learned from these messages showed uncertainty patterns that sensitive to the mode of communication, and were similar to the uncertainty patterns of their teachers.

Keywords: Social Learning; Rational Pedagogy; Rational Speech Act; Bayesian Inference

Introduction

When teaching concepts to others, there are two main approaches we can take. The first is to provide examples of objects that do or do not belong to the concept, relying on the learner's ability to generalize from examples. The second approach is to define the concept explicitly by specifying the features that are relevant for its definition and outlining its boundaries. To date, most models and theories of concept learning in cognitive science have been concerned with the first form of learning, with classical experimental paradigms involving children or adults learning concepts from examples (Rosch & Mervis, 1975; Shepard et al., 1961; Tenenbaum, 1999). Similarly, recent computational theories of pedagogy have focused on how teachers choose examples to provide to learners (Chen et al., 2024; Shafto et al., 2014; Vélez et al., 2023), rather than on cases when teachers directly communicate a defining rule. Nevertheless, humans frequently learn and teach concepts through explicit rules: we write dictionaries and legal statutes that carefully define concepts for others to interpret, and our formal education systems often ask students to recite explicit definitions related to the subject matter.

Many factors might lead a teacher to rely on examples rather than explicit definitions: the formal definition of a concept may not align with how the knowledge is intuitively perceived and represented (imagine teaching the color "red" to a child by telling them that red light has a wavelength of 625–750 nanometers), or the language may lack some words necessary to communicate rules effectively (Zettersten et al., 2024). Such considerations, however, are not the focus of the present research. Instead, the model, and the empirical study that follows, demonstrate that even when examples and rules are expressed using the same representational system (i.e., numbers), teachers' preferences for each mode of transmission additionally depend on their *uncertainty* about the concept boundaries.

We often use, and even teach to others, concepts whose boundaries we regard as uncertain. We might not be certain at what age a person becomes "old", but would nevertheless confidently cite our grandparents as examples if asked about the meaning of the word. In such scenarios, giving examples can help the learner develop a probabilistic, graded representation of the concept "old" – similar to the one we, as teachers, possess – that will allow them to apply the concept flexibly in different contexts. In contrast, communicating your best guess in the form of a clear-cut rule – *"every person above 70 years of age is old"* – might be counterproductive, as it will make the learner overconfident in a hypothesis you, as a teacher, are uncertain about.

In this project, we formalize this notion, deriving a normative model of when a rational teacher should choose to communicate a concept via examples vs. rules. Our analysis is built on the assumption that the objective of teaching is not only to convey rules, but also the graded uncertainty structure regarding the nature of these rules (stated formally, to communicate $p(H)$ instead of h^*). Of course, both clear-cut rules and graded concept representations can be communicated in both ways. The *size principle* in Bayesian concept learning (Tenenbaum, 1999, 1998) guarantees that after being given many examples, rational learners will represent the concept with clear-cut boundaries. In contrast, a probabilistic, graded concept representation can be communicated in the form of a rule by laboriously stating the level of confidence for different situations ("Be 100% sure that a person over 80 is old, 90% sure about 70 to 80…", etc.). However, in accordance with the rational speech act theoretical framework (Frank & Goodman, 2012; Hawkins et al., 2023), we model teachers as being *cost-sensitive*, namely, trying to balance the message's informative value with its length. As we shall see, this will lead them to communicate high-uncertainty concepts through examples and low-uncertainty ones through definitions.

In G. Boleda, I. Dautriche, M. Maldonado, C. Saldana and T. Torrent (Eds.), *Proceedings of the 48th Annual Conference of the Cognitive Science Society* ©2026 the author(s). This work is licensed under a Creative Commons Attribution 4.0 International License (CC BY).

Computational Model

Settings

The model assumes a continuous concept learning task where the goal is to learn a mapping from a point in R^k (a location in a k-dimensional feature space) to a binary category label. The hypothesis space consists of concept definitions that correspond to continuous regions of the feature space and can be represented as a lower and upper bound for all dimensions. Each hypothesis maps a coordinate to concept membership:

$$h_{\mathbf{b}^0,\mathbf{b}^1}(\mathbf{x}) = \begin{cases} 1, & \text{if } b_i^0 \leq x_i \leq b_i^1 \text{ for all } i \in \{1,\ldots,k\}, \\ 0, & \text{otherwise.} \end{cases}$$

In the case of a unidimensional feature space (where our simulations and experiment are set), hypotheses are simply ranges of numbers, defined by an upper and a lower bound.

A learner is provided with a message m (either a rule or a set of examples $x^1,\ldots,x^k$) and uses hypothesis averaging to decide, for each new object encountered, its probability of belonging to the concept (posterior predictive distribution).

$$p(x \in C) = \int h(x)p(h \mid m)dh$$

Learning Process

Learning From Examples When the learner receives a set of k positive examples $X = x^1,\ldots,x^k$ as their message, they use a previously established Bayesian inference algorithm (Tenenbaum, 1998, 1999) to update their posterior over the hypothesis space. Specifically,

$$p(h \mid X) \propto p(X \mid h)p(h)$$

where the likelihood $p(X \mid h)$ is given by the size principle under the strong sampling assumption:

$$p(X \mid h) = \begin{cases} \left(\frac{1}{|h|}\right)^k, & \text{if } h(x^i) = 1 \ \forall i \\ 0, & \text{otherwise} \end{cases}$$

Learning From Rules Alternatively, the learner can receive a message in the form of a rule, rather than data consistent with it. Such a message can be represented as lower and upper bounds for all the m dimensions – exactly how hypotheses are represented. Assuming a perfect communication channel without any noise added, the learner updates their posterior in the following way when provided with a rule R:

$$p(h \mid R) = \begin{cases} 1, & \text{if } R = h \\ 0, & \text{otherwise} \end{cases}$$

Teaching Process

Our model draws inspiration from previous work on rational pedagogy (Shafto et al., 2014), which defines the teacher objective as increasing the learner's belief in the true hypothesis about the target concept. However, we add a crucial component to this model: the teacher can have *uncertainty* over what the correct hypothesis is. Thus, the objective of the teacher is not to increase the learner's belief in a single hypothesis, but rather to replicate their own structure of uncertainty (posterior over H) in the learner's mind. We use the Kullback-Leibler (KL) divergence as a measure of the correspondence between the teacher's and learner's posterior distributions, and set the objective of the teacher to find the message that minimizes it.

$$m^* = \arg\min_m D_{\mathcal{KL}}\left(P^{\text{teacher}}(H) \,\middle\|\, P^{\text{learner}}(H \mid m)\right)$$

However, our teachers are *cost-sensitive* communicators. They do not only care about sending the message that would reflect their hypothesis distribution with the highest precision, but also prefer to avoid sending long messages. Thus, we define the message utility function to be:

$$U(m) = -D_{\mathcal{KL}}\left(P^{\text{teacher}}(H) \,\middle\|\, P^{\text{learner}}(H \mid m)\right) - \lambda * \text{Cost}(m)$$

where λ is a free cost-sensitivity parameter, and the cost of a message is determined by the number of numbers included in the message. For example, in the unidimensional case, the cost of sending a rule is 2 (because the teacher needs to specify one lower bound and one upper bound), and the cost of sending examples is just the number of examples included in the message.

Teaching With Examples When teaching with examples, the objective of the teacher is to find the set X of k positive examples with the highest utility according to the utility function specified above. However, if we allow large example sets, the number of possible example sets explodes combinatorially (there are about 2×10^{13} example sets of up to 10 examples). Therefore, we use a greedy algorithm, where at each step, the teacher chooses between either adding the single example that maximizes the message utility or stopping and sending the already curated example set.

$$X_{t+1} = \begin{cases} X_t \cup \arg\max_x U(X_t \cup \{x\}) & \text{if } U(X_t \cup \{x\}) > U(X_t), \\ X_t & \text{otherwise.} \end{cases}$$

Teaching With Rules When teaching with rules, the teacher chooses the rule that maximizes the utility function described above: $r^* = \arg\max_r U(r)$.

Deciding Between Examples and Rules The teacher decides whether to deliver a rule or a set of examples based on their respective utilities in a stochastic way using a softmax:

$$P(\text{choose } r \mid U_r, U_X) \propto \exp(\beta U_r)$$

where β is an inverse temperature parameter.

Simulations

Simulation code, as well as experimental data and analysis code, are available on the project GitHub repository:

https://github.com/arielevy8/examples-rules-uncertainty. We explore the model through simulations in a unidimensional setting, where the target concept is a continuous range of integers between 1 and 100. This setting is a simplified version of the number-game paradigm used in previous research (Ellis, 2023; Tenenbaum, 1998), in which concepts are defined solely by numerical ranges rather than by mathematical properties or other rules[1].

In the simulations, we examined the utility of teaching with rules and examples across varying levels of the teacher's cost sensitivity and uncertainty about the target concept. To manipulate teachers' uncertainty, teachers first learned the concept from examples before teaching it to others using examples and rules. Cost sensitivity values ranged from 0 to 1.5, and the number of examples observed by the teacher ranged from 5 to 60. For each parameter combination, we ran 15 simulation repetitions, with the ground-truth concept randomly initialized in each run.

Figure 1A exemplifies the simulations in two specific scenarios. In the left panel, the teacher has learned the target concept with only a few examples, reflected in the posterior predictive distribution plotted. They can communicate the concept via three examples, resulting in a virtually identical posterior distribution for the learner ($D_{\mathcal{KL}} \approx 0$), or communicate a rule, resulting in a similar maximum a posteriori (MAP), but a very dissimilar posterior distribution. The teacher in this case would rather pay the tiny added cost (3 numbers in the example set vs. 2 numbers in a definition) to effectively communicate their uncertainty structure. In contrast, if the teacher is very confident regarding the boundaries of the concept (e.g., if they have seen 40 examples; right panel), communicating a rule would lead to a similar posterior in the learner's mind, but achieving a similar posterior with examples will require many of them. The teacher, being cost-sensitive, will send the rule.

The utility of rule-based and example-based messages, as a function of teacher conceptual certainty, is plotted in Figure 1B. As expected, the utility of using examples decreases as certainty increases, while the utility of using rules increases with certainty. The left and right panels demonstrate this effect for different degrees of cost sensitivity (λ) values—which do not change the main trend, but rather the point at which the teacher will prefer to switch from examples to rules.

Study 1

Our simulations support a clear prediction, that can be tested empirically: rational teachers would be more likely to communicate concepts with rules when they are highly certain about the concept boundaries. We designed an experiment to test whether prolific participants, assigned the role of teachers in a range-based concept learning task, act according to this prediction.

[1]Limiting the hypothesis space to numerical ranges allows us to have a straightforward definition of the cost of providing a rule—unlike rules based on mathematical properties, for which utterance costs are more ambiguous.

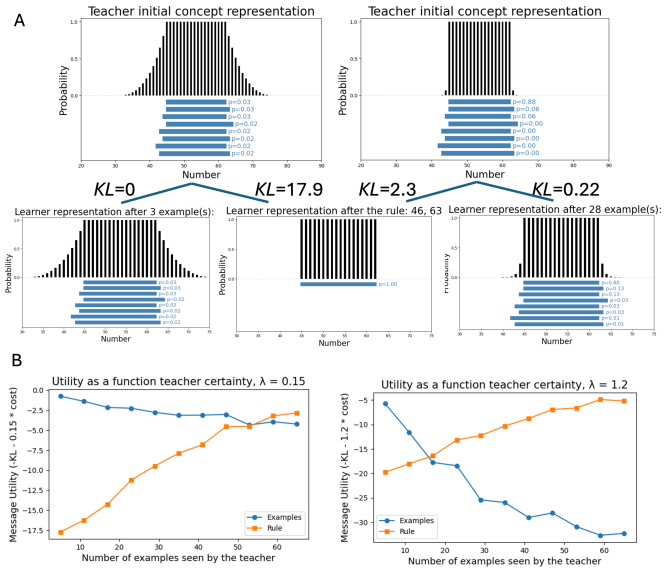

Figure 1: Simulations results. *A*) Examples of the teaching outcomes of an uncertain teacher (top left) and a certain teacher (top right). Black vertical lines represent posterior predictive distribution. Blue segments represent top 8 underlying hypotheses, sorted by their posterior probability. *B*) Message utility as a function of teacher's uncertainty across low (left) and high (right) cost sensitivity parameters.

Methods

Sample size, as well as experiment details and analysis plan, were pre-registered (https://aspredicted.org/2rc4vy.pdf). One hundred US-based Prolific participants were recruited for this study. Participants took part in a concept learning and teaching game, where the concepts to be learned are specific ranges of numbers between 1 and 100. At the beginning of the study, participants are introduced to the background story of the game: they are assigned the role of spacecraft engineers on a remote planet. Each spacecraft has a specific range of safe temperatures in which it can operate, measured in a unique temperature unit that ranges from 1 to 100. For each spacecraft, participants complete the following three phases:

Phase I – active learning In this phase, the participant interacts with the number line, testing different temperatures and receiving feedback on whether these temperatures belong to the safe range of the current spacecraft. The *key experimental manipulation* occurs in this phase: after finding the first safe temperatures, the number of additional interactions with the number line is limited. In the *low certainty* condition, participants can only test two additional temperatures, making it impossible to understand the safe range precisely. In the *medium certainty* condition, they are allowed 6 additional tests. In the *high certainty* condition, they are allowed 12 additional tests, making sampling positive and negative

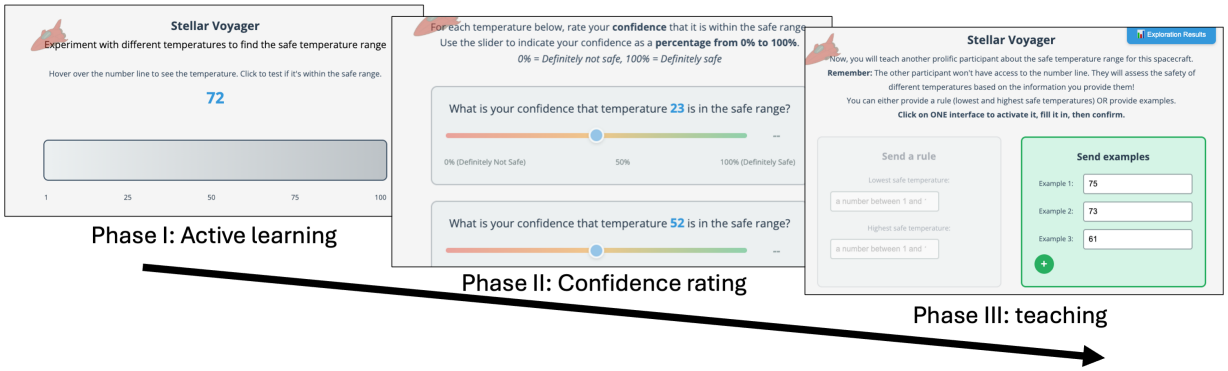

Figure 2: A depiction of a trial sequence. In each trial, participants completed three phases: active learning, confidence ratings, and teaching.

examples around the boundary and inferring a clear-cut rule-based concept very likely.

Phase II – posterior predictive report We asked participants to report, for 15 different temperatures between 1 and 100, their degree of confidence that each temperature belongs to the concept (the safe range). This phase served as a way to sample directly from the posterior predictive distribution of participants and allowed us to infer their uncertainty level regarding the concept—both for the purpose of a manipulation check and to see if the uncertainty level predicted their frequency of using examples vs. rules. The 15 temperatures they are asked about are sampled stochastically from the number line, and numbers that are closer to the ground-truth concept were more likely to be asked about.

Phase III – teaching In this phase, participants were told they could teach an additional Prolific participant—who would not have the opportunity to interact with the number line—about the safe temperature range for the current spacecraft. They are presented with two user interfaces, side by side: one that allows them to deliver a rule (a lower bound and an upper bound), and one that allows them to deliver examples of safe temperatures. They can add additional examples using a "+" button, without limitation on the number of examples. They can interact with both UIs, but can only choose one type of message to send.

Participants completed all three phases for 6 spacecraft—two in each certainty condition. After completing the study, participants reported basic demographics, completed an attention check, and were debriefed.

Results

Manipulation check We used an inverse inference algorithm to recover the hypothesis distribution most likely to generate teachers' Phase II confidence ratings and computed its Shannon entropy on each trial. As shown in Figure 3A, entropy decreased as participants were given more time to interact with the number line. A mixed-effects regression confirmed a significant effect of condition on entropy: compared to the low condition, entropy was lower in the medium condition ($\beta = -0.65$, $SE = 0.09$, $t = -7.00$, $p < .001$) and lowest in the high condition ($\beta = -1.06$, $SE = 0.09$, $t = -11.36$, $p < .001$).

Uncertainty leads to using examples To test the main hypothesis—that participants would prefer rule-based messages when they were more certain about the concept boundaries—we fit a generalized linear mixed-effects model predicting teaching method from condition, with random intercepts for participants. The model revealed a significant effect of condition on the use of rules ($p < .001$): compared to the low condition, participants in the medium condition showed higher odds of rule-based teaching ($\beta = 1.10$, $SE = 0.31$, $z = 3.48$, $p < .001$), with an even stronger effect in the high condition ($\beta = 2.33$, $SE = 0.37$, $z = 6.23$, $p < .001$). Figure 3B illustrates this pattern by showing the proportion of teaching modes across conditions.

As a complementary analysis, we examined teaching choices as a function of trial-by-trial hypothesis entropy. Consistent with the condition-based results, participants were more likely to use examples as entropy increased. A generalized linear mixed-effects model with random intercepts for participants confirmed this relationship, showing that higher entropy predicted greater odds of example-based teaching ($\beta = 0.90$, $SE = 0.17$, $z = 5.43$, $p < .001$). Figure 3C plots the predicted probability of using examples as a function of teacher entropy.

Properties of examples given As an exploratory analysis, we examined which examples participants provided when opting for example-based teaching. The vast majority of examples (87%) fell within the participant's *known safe range*—values that could be classified as belonging to the target concept with certainty based on Phase I (exploration). As shown in Figure 3D, participants frequently selected examples near the boundaries of this range. Both the restriction to highly certain examples (with posterior predictive probability of 1) and the preference for boundary cases mirror qualitative patterns observed in our computational simulations.

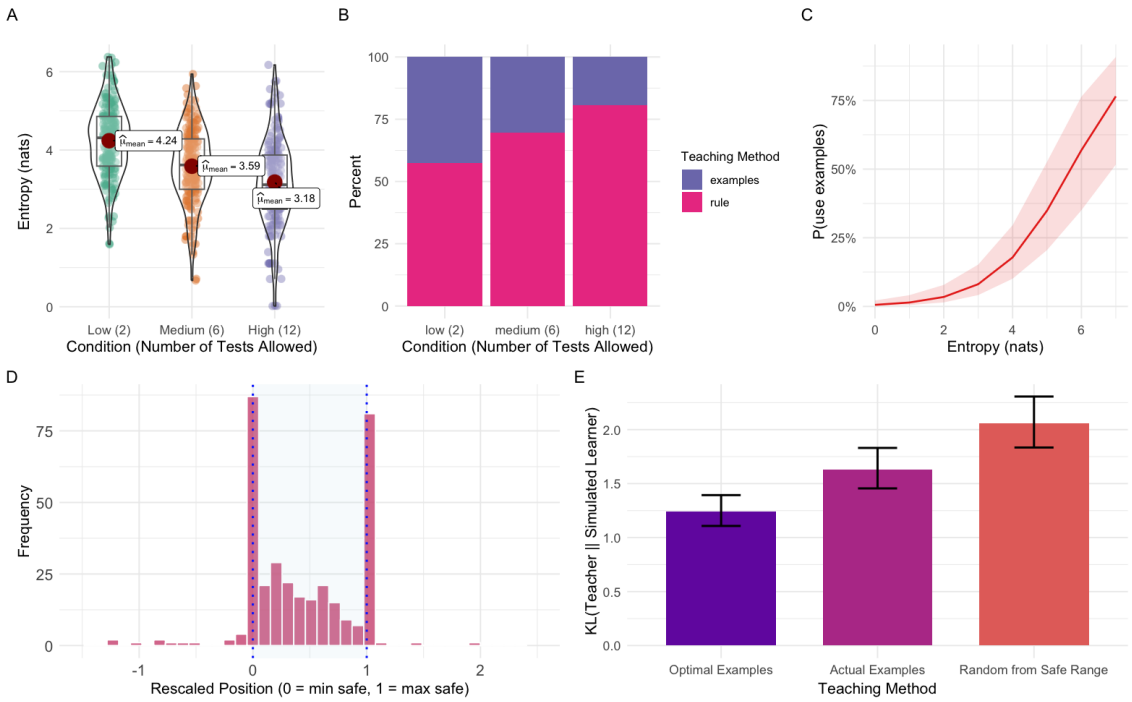

Figure 3: Study 1 results. A) Trial-level entropy (recovered from participants' confidence ratings) as a function of experimental condition. B) Teaching method as a function of experimental condition. C) Teaching method as a function of entropy (effect plot of the logistic model). D) Histogram of the examples given, rescaled relative to the *known safe range* (max and min temperatures that were observed as safe during exploration). E) KL-divergence between the teachers' (recovered) posterior distributions and those of simulated learners receiving random, actual, and optimal examples. Error bars represent 95% CIs.

To test whether teachers' examples were effective in communicating their underlying uncertainty structure, we computed the KL divergence between each teacher's inferred posterior and a simulated rational learner exposed to their examples. We compared the KL divergence produced by teachers' actual examples to two baselines: optimal examples generated by the greedy algorithm outlined above, and examples sampled randomly from the safe range. We assigned the KL values to a generalized linear mixed-effects model with a Gamma distribution and log link and random intercepts for participants.[2] We found that actual examples reduced KL divergence ($\beta = -0.23$, $SE = 0.08$, $p = .004$), while optimal examples showed an even larger reduction ($\beta = -0.50$, $SE = 0.08$, $p < .001$), suggesting that participants' example choices were systematically more informative about their uncertainty structure than random ones, though they were far from being fully optimal (Figure 3E).

Study 2

In Study 2, we recruited new participants to act as learners receiving messages produced by teachers in Study 1, while being informed about how those messages were generated. Study 2 had two goals: first, to test whether learners who received examples developed greater uncertainty about the target concept than those who received rules; and second, to assess whether the examples provided in Study 1 were effective at conveying teachers' uncertainty to real learners.

Methods

Participants and design One hundred US-based participants were recruited via Prolific. All participants completed four within-participant learning conditions (one trial per condition) in a 2×2 design. The first factor was teaching mode, which varied whether the message consisted of example sets or explicit rules. The second factor was message source, which varied whether the message was given by a real participant from Study 1 or generated at random based on an actual concept presented in Study 1.

Random example sets were generated by first identifying trials in which teachers provided examples, then sampling examples uniformly at random from the known safe range based on the teacher's exploration in these trials. The number of examples in each random set was matched to teachers' behavior by sampling with noise around the average number of examples they provided.

Procedure The instructions introduced participants to the generating process of the examples or rules they were about

[2] A Gamma model was used for KL values here and in the following study because these values are bounded at 0 and were right-skewed; results are robust to Gaussian specifications.

to see. Specifically, they were informed that their teachers were allowed to test temperatures to learn about the spaceship's temperature range, and that in some cases they had more time to experiment than others. In addition, they were informed that teachers could choose between message types. Next, they were shown the message given by previous participants. Then, they responded to the same posterior predictive report interface that was used in Phase II of Study 1. The temperatures for which they reported concept membership confidence were the exact same ones for which the corresponding teacher responded, to allow a clean comparison between the representations of the teacher–learner pairs.

Results

Uncertainty is higher when learning from examples We used an inverse inference algorithm to recover the posterior distribution most likely to generate teachers' reported confidence ratings and computed its Shannon entropy on each trial. A linear mixed-effects regression with message type and message source as fixed effects, and random intercepts for participant ID, revealed a significant main effect of message type: entropy was lower for rule-based messages than for example-based messages ($\beta = -0.52$, $SE = 0.11$, $t = -4.91$, $p < .001$, Cohen's $d = 0.37$). In contrast, there was no significant effect of information source (actual vs. random: $\beta = -0.11$, $SE = 0.10$, $t = -1.09$, $p = .28$).

Actual examples transmit uncertainty structure better than random ones To test whether teachers' examples preserve uncertainty structure for real learners, we computed the KL divergence between teachers' inferred posteriors and the posteriors of learners exposed to those examples. We compared learners who received teachers' actual examples to learners who received example sets sampled randomly from the known safe range. A generalized linear mixed-effects model revealed a significant effect of source: learners exposed to random examples showed higher KL divergence than those who learned from teachers' actual examples ($\beta = 0.13$, $SE = 0.06$, $p = .027$, Cohen's $d = 0.28$). Consistent with our model predictions, teachers selected examples that conveyed their uncertainty structure more effectively than would be expected from random choices within their known safe range.

Discussion

In this paper, we drew on principles from Bayesian concept learning (Tenenbaum, 1999) and rational pedagogy (Shafto et al., 2014) to develop a theory of how teachers choose between rule-based and example-based concept teaching. A central prediction of the model, established through simulations, is that teachers should prefer examples when they are uncertain about a concept's boundaries, because examples allow them to efficiently recreate their uncertainty structure in the learner's mind. This prediction was supported in Study 1: participants who actively learned number-range concepts subsequently adjusted their teaching strategies as a function of their own uncertainty, in the direction predicted by the model. When new participants learned from these messages (Study 2), teachers' rules induced more clear-cut concept representations, whereas example sets were found to transmit the teachers' uncertainty structure faithfully.

Previous work on rational pedagogy (Shafto et al., 2014; Vélez et al., 2023) has focused on settings in which the teacher knows a rule for certain but can communicate it only through examples. In these cases, learners can often infer a precise rule from a small number of examples by reasoning about what led the teacher to choose the specific examples they did. This line of work might suggest that learning from pedagogical examples (unlike non-pedagogical example sampling) generally leads to confident learners (Bonawitz et al., 2011). In contrast, our results from Study 2 show that learners who received examples develop a concept representation with high uncertainty. This apparent tension can be reconciled by noting that pedagogical reasoning depends on learners' beliefs about the teacher's epistemic state. When learners consider the possibility that the teacher might also be uncertain, and especially when they know that a teacher could have provided a rule but instead chose to provide examples, learners infer greater uncertainty about the target concept.

This study focused on simple, unidimensional concepts defined over numerical ranges. When studying more complex concepts, cognitive scientists distinguish between rule-based categorization, where membership is determined by applying an explicit, all-or-none rule, and family-resemblance categorization (Rosch & Mervis, 1975), where some items are perceived as more representative (or prototypical) of the concept than others. The Bayesian approach to concept learning (Tenenbaum, 1999), which we adopt, unifies these: learners consider a space of candidate *rules* (hypotheses), and representativeness effects arise because some items are consistent with more plausible hypotheses (Tenenbaum & Griffiths, 2001). On this view, graded membership is a marker of conceptual uncertainty. Thus, our model predicts that concepts that elicit family-resemblance judgments, like power (Haugaard, 2010), will be communicated using examples, whereas concepts that elicit all-or-none judgments will be communicated by directly conveying the rule. Future work is needed to test this prediction with more natural concepts.

While we believe the uncertainty factor identified here to be an important determinant of the choice between rules and examples, it is certainly not the only one. In many real-world teaching contexts, even low-uncertainty concepts are best communicated using a combination of explicit rules and illustrative examples, particularly when rules are abstract or costly to fully specify. In such cases, examples can ground otherwise precise but underspecified rules by conveying concrete information that would be inefficient to communicate explicitly. Extending the present framework to allow teachers to combine rule-based teaching with examples remains an important direction for future work.

Acknowledgments

We thank all members of the Laboratory for Social Cognitive Science at Harvard for their valuable feedback throughout the development of this project. We acknowledge the use of large language models (specifically the Claude Opus family and ChatGPT) for assistance with code generation and copyediting. All code was carefully reviewed and validated by the authors, who take full responsibility for its correctness. This work was supported by award number N00014-22-1-2205 to FC from the Office of Naval Research.

References

Bonawitz, E., Shafto, P., Gweon, H., Goodman, N. D., Spelke, E., & Schulz, L. (2011). The double-edged sword of pedagogy: Instruction limits spontaneous exploration and discovery. *Cognition*, *120*(3), 322–330. doi: https://doi.org/10.1016/j.cognition.2010.10.001

Chen, A. M., Palacci, A., Vélez, N., Hawkins, R. D., & Gershman, S. J. (2024). A Hierarchical Bayesian Model of Adaptive Teaching. *Cognitive Science*, *48*(7), e13477. doi: https://doi.org/10.1111/cogs.13477

Ellis, K. (2023). Human-like Few-Shot Learning via Bayesian Reasoning over Natural Language. *Advances in Neural Information Processing Systems*, *36*, 13149–13178.

Frank, M. C., & Goodman, N. D. (2012). Predicting Pragmatic Reasoning in Language Games. *Science*, *336*(6084), 998–998. doi: https://doi.org/10.1126/science.1218633

Haugaard, M. (2010). Power: A 'family resemblance' concept. *European Journal of Cultural Studies*, *13*(4), 419–438. doi: https://doi.org/10.1177/1367549410377152

Hawkins, R. D., Franke, M., Frank, M. C., Goldberg, A. E., Smith, K., Griffiths, T. L., & Goodman, N. D. (2023). From partners to populations: A hierarchical Bayesian account of coordination and convention. *Psychological Review*, *130*(4), 977.

Rosch, E., & Mervis, C. B. (1975). Family resemblances: Studies in the internal structure of categories. *Cognitive Psychology*, *7*(4), 573–605. doi: https://doi.org/10.1016/0010-0285(75)90024-9

Shafto, P., Goodman, N. D., & Griffiths, T. L. (2014). A rational account of pedagogical reasoning: Teaching by, and learning from, examples. *Cognitive Psychology*, *71*, 55–89. doi: https://doi.org/10.1016/j.cogpsych.2013.12.004

Shepard, R. N., Hovland, C. I., & Jenkins, H. M. (1961). Learning and memorization of classifications. *Psychological Monographs: General and Applied*, *75*(13), 1–42. doi: https://doi.org/10.1037/h0093825

Tenenbaum, J. B. (1998). Bayesian Modeling of Human Concept Learning. In *Advances in Neural Information Processing Systems* (Vol. 11). MIT Press.

Tenenbaum, J. B. (1999). Rules and Similarity in Concept Learning. In *Advances in Neural Information Processing Systems* (Vol. 12). MIT Press.

Tenenbaum, J. B., & Griffiths, T. L. (2001). The rational basis of representativeness. In *Proceedings of the 23rd annual conference of the Cognitive Science Society* (Vol. 6).

Vélez, N., Chen, A. M., Burke, T., Cushman, F. A., & Gershman, S. J. (2023). Teachers recruit mentalizing regions to represent learners' beliefs. *Proceedings of the National Academy of Sciences*, *120*(22), e2215015120. doi: https://doi.org/10.1073/pnas.2215015120

Zettersten, M., Bredemann, C., Kaul, M., Ellis, K., Vlach, H. A., Kirkorian, H., & Lupyan, G. (2024). Nameability supports rule-based category learning in children and adults. *Child Development*, *95*(2), 497–514. doi: https://doi.org/10.1111/cdev.14008