Metaphors we teach by: Uncovering the structure of metaphorical lay theories of teaching

Bridgette Martin Hard [a]*, Nathan Liang [a], Michelle Wong [a], and Stephen J. Flusberg[b]

aDepartment of Psychology & Neuroscience, Duke University, Durham, NC, USA

bDepartment of Psychology, SUNY Purchase College, Purchase, NY

*corresponding author: Bridgette Martin Hard, Department of Psychology & Neuroscience, Duke University, 417 Chapel Drive, Box 90086, Durham, NC, 27708 Email: bridgette.hard@duke.edu

Abstract

Teaching is a complex activity that people often discuss metaphorically, as when a professor is described as a *sculptor molding impressionable* students. What do such metaphors reveal about how people conceptualize teaching? Previous work has addressed related questions largely via researcher intuition and qualitative analyses of teacher attitudes. We sought to develop a more principled method for mapping the entailments of metaphorical concepts, using teaching metaphors as a case study. We presented participants with one of four common metaphors for the teacher-student relationship (identified in a preliminary study) and asked them to rate the degree to which a series of teacher attributes fit the metaphor. We then used iterated exploratory factor analysis to identify a small number of dimensions that underlie people's conceptions of teachers and examined whether the metaphors systematically differed along these dimensions. We found that teaching metaphors bring to mind distinct, coherent clusters of teacher attributes and different intuitions about teacher responsibility and power—a finding we replicated in a larger, pre-registered follow-up study using a new set of participants. This work provides a novel method for mapping the entailments of metaphorical concepts and sets the stage for educational interventions centered on shifting lay theories of teaching.

Keywords: metaphors, lay theories, concepts, teaching

1. Introduction

According to psychologists, people are _amateur scientists_ who actively explore the environment and develop intuitive, lay theories for how the world works (Gopnik, Meltzoff, & Kuhl, 1999). In turn, these lay theories shape how people understand and respond to new events and situations, guiding thought and action. For example, many people conceptualize the economy as a zero-sum system—a _pie_ where everyone must fight for their _slice_. White Americans who hold this view have been shown to feel especially threatened when presented with data about impending demographic changes in the US, leading them to endorse more conservative policy positions compared to those who hold a non-zero-sum economic theory (Perkins, Dils, & Flusberg, _under review_).

As we have just illustrated, one way that intuitive beliefs are conveyed is via metaphor. Metaphors allow people to draw on familiar, common knowledge of a basic source domain (_science, pies_) to talk about a more complex or abstract target domain (_learning, economics_). Over and above their role in communication, however, scholars have argued that people actually _think_ metaphorically about complex and abstract issues (Gibbs, 1994; Lakoff & Johnson, 1980; Thibodeau, Matlock, & Flusberg, 2019).

1. 1. Metaphors in Education

One domain where metaphorical lay theories have been shown to have significant real-world consequences is in education (e.g., Blackwell, Krzesniewski, & Dweck, 2007; Landau, Oyserman, Keefer, & Smith, 2014; Low, 2008). Research over the last two decades, for example, suggests that students' beliefs about the malleability of intelligence (also known as "mindsets") drive educational achievement (Yeager & Dweck, 2012). Those who consider intelligence to be a trait that can improve through hard work—the BRAIN IS A MUSCLE metaphor—react better to

performance setbacks than those who view intelligence as a fixed entity, resulting in superior long-term learning outcomes. However, interventions aimed at changing these lay theories to improve student performance have generated inconsistent results (e.g., Sisk et al., 2018; Yeager et al., 2019). One reason for this inconsistency may be that academic performance is influenced by numerous lay theories, not just about intelligence. Because learning occurs in a social context that includes teachers and the broader classroom environment, the intuitive beliefs students hold about the nature of teaching itself might also impact academic behaviors and outcomes (Low, 2015; Seung, Park, & Jung, 2015; Wan, Low, & Li, 2011).

Like other complex activities, people frequently talk about teaching metaphorically. When we describe a teacher as '*molding impressionable* students,' for example, we imply that the teacher is like a *sculptor* and students are like *clay*. The TEACHER AS A SCULPTOR metaphor may reflect the intuitive belief that learning is passive and that the teacher—and not the student––largely determines learning outcomes. To date, metaphors for teaching have largely been explored in the teacher education literature by scholars reflecting on their own practices or through a qualitative analyses of teacher attitudes (e.g., Alger, 2006; Mahlios, Massengill-Shaw, & Barry, 2010; Martıńez et al., 2001; Saban et al., 2007; Seferoğlu, Korkmazgil, & Ölçü, 2009; Shaw, Barry, & Mahlios, 2008; Tobin, 1990; but see Low, 2015 for a more rigorous analytical framework). This research suggests that teachers endorse a variety of different metaphors for thinking about their role as teachers (e.g. Chen, 2003). Notably, however, most prior work does not address the metaphors that *students* hold for teaching, nor the consequences of holding these metaphors (but see Wan, Low, & Li, 2011).

1. 2. Mapping Metaphorical Entailments

How can we best understand students' metaphorical lay theories of teaching and the extent to which they influence thought and action? An important first step is to identify which metaphors people actually endorse and to map out the *entailments* of these metaphors; that is, the associated ideas and inferences licensed by the metaphorical comparison (Low, 2015). For example, one potential entailment of the *BRAIN IS A MUSCLE* metaphor described earlier is that, while hard work may increase someone's intelligence or abilities, working too long on any one task may become cognitively exhausting, much like continuous physical exertion tires out muscles. These entailments, it is argued, provide insight into the mental model people use to represent the target domain and allow for empirically-informed predictions about how metaphors reflect and shape thinking (Low, 2015; Thibodeau, Crow, & Flusberg, 2017).

The most common approach to mapping entailments is to examine the figurative language people use in everyday speech and apply a commonsense understanding of the observed source domains—as we have demonstrated in the preceding paragraphs. Cognitive linguists have used this approach to isolate the structural schemas that underlie many concepts, from emotion to politics to time, across a variety of languages and cultures (e.g., Kövecses, 2005; Lakoff and Johnson, 1980). Much of the research cited previously about teaching metaphors employed this method. However, there are several issues with relying on intuition and patterns of language alone to make inferences about conceptual representations (Murphy, 1996). For example, the meaning of common metaphorical expressions might seem obvious and intuitive but could also reflect a post-hoc rationalization based on one's preexisting understanding of the expression (Keysar & Bly, 1995).

Another approach to mapping entailments is to ask a set of naïve participants to freely generate the entailments of a source domain in order to derive a set of conceptually coherent

metaphorical entailments for the target domain. For example, for the metaphor 'crime is a *virus*,' participants might be asked how they would solve a literal virus problem in their city. The responses would then be used to predict which solutions to a city's crime problem would be conceptually congruent with the CRIME IS A VIRUS metaphor (Thibodeau & Boroditsky, 2011). To validate these experimenter intuitions, another set of participants might be asked to match specific solutions to a city's crime problem to specific metaphors, either a CRIME IS A VIRUS or CRIME IS A BEAST metaphor (Thibodeau & Boroditsky, 2013; See also Davis, Watson, & Bakerson, 2015 for a different crowdsourcing approach to mapping the entailments of metaphors for the role of school principals).

1. 3. The Present Research

In this paper, we aimed to develop a novel systematic method for mapping the structural entailments of a complex metaphor, focusing specifically on metaphors for teaching. Our approach was inspired in part by traditional psychometric methods that have been used to uncover the dimensional structure of personality traits and other psychological constructs. Personality researchers have been using factor analysis for nearly a century to take participant ratings of different, but interrelated attributes and reduce them to a smaller set of latent factors, which constitute the personality traits of a given theory (e.g. Flanagan, 1935; Lee & Ashton, 2007). As we illustrate, a similar approach can be used to reduce ratings of attributes associated with a specific metaphor down to a smaller set of latent factors that constitute the conceptual entailments, or '*personality traits*' of the metaphor.

In the present studies, we focused on four metaphors for the teacher-student relationship that our preliminary research—described in the Online Supplement—identified as relatively common and apt among current and former college students: TEACHER AS: GARDENER,

COACH, TOUR GUIDE, or SCULPTOR. In Study 1, we presented participants with a list of statements describing teacher attributes alongside one of the four metaphors. Participants rated the degree to which each attribute of college teachers fit with the assigned metaphor and rated each metaphor on a few additional features. We used exploratory factor analysis to uncover a smaller set of latent factors underlying the larger collection of teacher attributes. This allowed us to identify a distinct, coherent cluster of teacher attributes associated with each metaphor, providing a principled way of mapping the structure of metaphorical lay theories. In Study 2, we replicated these findings with a new, larger sample and a preregistered analysis. This work demonstrates a novel method for systematically mapping the entailments of metaphorical lay theories and provides insight into the metaphors for college teaching endorsed by many Americans today.

2. Study 1

The primary goal of Study 1 was to identify and map out the conceptual entailments associated with each of four metaphors for the teacher-student relationship, which we assessed for popularity and aptness in a preliminary study (see Flusberg & Hard, 2019, as well as the Online Supplement):

- GARDENER: *"A teacher is like a gardener and students are like plants"*
- COACH: *"A teacher is like a coach and students are like athletes"*
- TOUR GUIDE: *"A teacher is like a tour guide and students are like tourists"*
- SCULPTOR: *"A teacher is like a sculptor and students are like clay"*

Our key question was whether different metaphors would be reliably associated with distinct clusters of teacher attributes, and whether this association could be measured in a systematic, principled way using factor analytic techniques. All measures used in the study, as well as

relevant data and analyses are available through the Open Science Framework at:

https://osf.io/e6vba/.

2.1. Participants

We recruited 399 participants to complete the survey through Amazon's Mechanical

Turk, collected at two different time periods[1]. We required that participants be a current or

former college student living in the U.S. or Canada, with an approval rating greater than 95% on

at least 100 prior Mechanical Turk tasks (known as Human Intelligence Tasks, or HITs).

Additionally, we narrowed our recruitment to participants between the ages of 18 and 30 in order

to sample participants close in age and experience to current college students. We excluded data

from two participants who did not enter a code to receive payment, one who reported an age

older than 30, one who did not report an age, three participants who provided duplicated (i.e.,

copy and paste) responses to all free response items, and another whose responses to free

response items indicated extremely weak proficiency in English. See Table 1 for demographic

data.

Table 1. Demographic characteristics of participants in Studies 1 and 2.

Demographic characteristic	Study 1 (N = 395)	Study 2 (N = 536)
Gender - Male	51%	46%
Race/Ethnicity		
White	64%	62%
Black	11%	11%
Asian	7%	12%
Hispanic/Latino	7%	6%
Amer.Indian/ Alaskan Native	< 1%	< 1%

Multiracial	8%	9%
College Status		
Currently enrolled	30%	48%
Graduated	57%	44%
Didn't graduate	14%	8%
Mean age (SD)	25.5 (3.0)	24.1 (2.8)

2.2. Procedure

Participants were randomly assigned to view one of the four teaching metaphors: *GARDENER*, *COACH*, *TOUR GUIDE*, or *SCULPTOR*. Participants were asked to think about what the metaphor means and jot down a sentence or two about how a college teacher could be like the metaphor that was described. Participants then viewed a list of 43 statements describing college teachers (e.g., '*Teachers transfer their knowledge to students*') and rated how well each item fit the metaphor that they were given (see Table 2). The statements were generated by consulting measures of teacher behavior (e.g., *The Teacher Behavior Checklist*, Keeley, Smith, & Buskist, 2006), the thematic analysis of free response data from our preliminary study (see Online Supplement), and the personal experience of the researchers. Participants were instructed to rate how well each item agreed with the metaphor they received on a scale of 1 ('not well at all') to 7 ('extremely well'). Participants were specifically asked to rate how well each item agreed with the metaphor, not whether they personally believed each item was true.

Next, participants answered four questions aimed at whether the metaphors captured different intuitions about teacher responsibility and power: They rated, according to the metaphor, (a) how much responsibility college teachers have for students' learning, (b) how

much responsibility students have for their own learning in college, (c) how much power college teachers have to influence what students learn, and (d) how much power college teachers have to influence how students develop as people. Each item was rated on a scale of 0-100, with 0 meaning "none at all" and 100 meaning "a great deal." Finally, participants viewed all four of the metaphors and selected their personal favorite.

2.3. Results

2.3.1. *Factor structure of teacher characteristics*

Our first step was to perform an exploratory factor analysis of the 43 different teacher attributes to identify the latent variables underlying the attributes and examine whether these variables differed across the metaphors.[2] Very simple structure and parallel analyses suggested five factors. We performed an initial principal axes factor analysis (oblimin rotation) in R, extracting five factors. Next, we worked iteratively to reduce the number of items and improve coherence by (1) examining the variables for normality and excluding especially skewed items, (2) removing items that loaded on multiple factors, and (3) examining the face validity of each item to identify and remove redundant items. Each step of this process is detailed on Open Science Framework.

Table 2. Pattern matrix of factors loadings for 20 teacher characteristics on the five factors in Study 2 and 3. Highest loading items for each factor (>.40) are shown, Factor loadings < .40 are shown in bold italics for teacher attributes that loaded strongly onto a factor in one study, but not in the other.

		Factor					
Teachers...		**1** **Nurturer**	**2** **Authority** **Figure**	**3** **Community-** **Builder**	**4** **Engaging** **Communicator**	**5** **Stimulator**	
Proportion of variance explained	S1	22%	21%	22%	19%	16%	
	S2	21%	21%	23%	17%	17%	**Uniqueness**

are sensitive to their students' needs	S1	0.85				0.32
	S2	0.81				0.34
adapt their teaching to different students' needs	S1	0.74				0.35
	S2	0.78				0.40
care about students' well-being	S1	0.73				0.41
	S2	0.69				0.37
get to know their students	S1	0.56				0.35
	S2	0.57				0.47
command respect	S1		0.76			0.26
	S2		0.72			0.36
establish classroom rules	S1		0.71			0.37
	S2		0.56			0.41
are authority figures	S1		0.67			0.49
	S2		0.77			0.44
have clear expectations for students	S1		0.60			0.48
	S2		0.67			0.44
encourage students to support one another	S1			0.73		0.22
	S2			0.74		0.26
encourage teamwork	S1			0.70		0.27
	S2			0.77		0.24
encourage students to get to know one another	S1			0.63		0.25
	S2			0.65		0.32
encourage a sense of community	S1			0.56		0.36
	S2			0.61		0.37
communicate clearly	S1				0.71	0.41
	S2				0.74	0.36
are confident	S1		*0.35*		0.61	0.32

	S2	0.49	*0.40*			0.45
are personable	S1		0.58			0.35
	S2		0.58			0.43
hold students' attention	S1		0.52			0.46
	S2		0.64			0.46
provoke debate	S1				0.73	0.33
	S2				0.83	0.23
promote class discussion	S1				0.65	0.26
	S2				0.52	0.35
stimulate students' thinking	S1				0.51	0.45
	S2				*0.26*	0.60
question students' ideas	S1				0.49	0.44
	S2				0.62	0.48

Note. 'Principal axis factoring' extraction method was used in combination with an 'oblimin' rotation

Our final factor analysis was conducted on 20 items. We extracted five factors using principal axis factoring, oblimin rotation. Four items loaded onto each factor. The pattern matrix is shown in Table 2, correlations between factors are reported in Table 3, and the correlation matrix is available on the Open Science Framework. Based on the highest-loading items, we interpreted the factors as:

(1) **Community-Builder** (e.g., "Teachers encourage students to support one another")

(2) **Nurturer** (e.g., "Teachers are sensitive to their students' needs")

(3) **Authority Figure** (e.g., "Teachers command respect")

(4) **Engaging Communicator** (e.g., "Teachers communicate clearly")

(5) **Stimulator** (e.g., "Teachers provoke debate")

Table 3. Factor correlations in Studies 1 and 2.

		1 Nurturer	2 Authority Figure	3 Community- Builder	4 Engaging Communicator	5 Stimulator
1 Nurturer	S1	—				
	S2	—				
2 Authority Figure	S1	0.27	—			
	S2	0.13	—			
3 Community- Builder	S1	0.52	0.49	—		
	S2	0.46	0.44	—		
4 Engaging Communicator	S1	0.47	0.55	0.49	—	
	S2	0.36	0.45	0.50	—	
5 Stimulator	S1	0.43	0.36	0.53	0.49	—
	S2	0.26	0.41	0.56	0.42	—

2.3.2. *How do teacher characteristics vary across metaphors?*

Once a 5-factor solution was applied, we derived and saved the factor scores using the regression method in R (Grice, 2001). Factor scores were only derived for participants with no missing values ($N = 381$). A multivariate ANOVA was performed on the five factor scores with metaphor (*GARDENER*, *COACH*, *TOUR GUIDE*, or *SCULPTOR*) as a between-subjects variable. Overall, metaphor condition showed a significant effect on the set of factors, $F(15, 1125) = 15.51$, $p < .001$, $\eta^2 = .17$. Univariate analyses of variance revealed significant effects of metaphor on each factor individually: the *nurturer* ($F(3,377) = 10.93$, $p < .001$, $\eta^2 = .08$), *authority figure* ($F(3,377) = 13.47$, $p < .01$, $\eta^2 = .10$), *community-builder* ($F(3,377) = 6.37$, $p < .001$, $\eta^2 = .05$), *engaging communicator* ($F(3,377) = 5.34$, $p < .01$, $\eta^2 = .04$), and *stimulator* factors ($F(3,377) = 3.20$, $p < .05$, $\eta^2 = .03$).
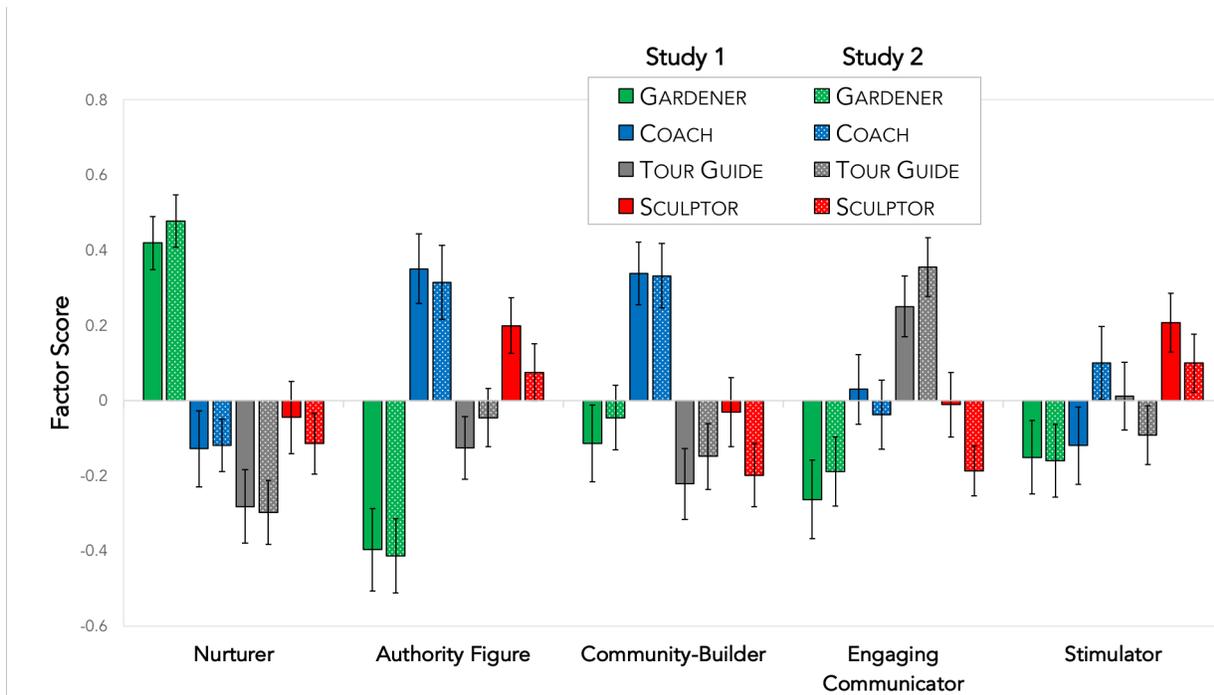
Figure 1. Mean factor scores for each of the five factors in Studies 1 and 2, as a function of the conceptual metaphor that participants considered. Study 1 results are shown in solid bars, Study 2 results in dotted bars. Error bars show +/- 1 SEM.

Figure 1 shows the pattern of differences across the metaphors for each factor. Visual examination of the figure shows striking differences in teacher characteristics across metaphors. The GARDENER metaphor entails that teachers are nurturers but not authority figures or engaging communicators. The COACH metaphor entails that teachers are authority figures and community-builders, but not necessarily nurturers or stimulators. The TOUR GUIDE metaphor entails that teachers are engaging communicators but not community-builders, nurturers, or authority figures. The SCULPTOR metaphor entails that teachers are authority figures and stimulators.

2.3.3. *Are different metaphors associated with different intuitions about responsibility?*

Metaphor had a small but reliable effect on ratings of how much responsibility college teachers had for students' learning, $F(3, 391) = 6.84$, $p < .001$, $\eta^2 = .05$ (see Figure 2). Post-hoc tests using bonferroni adjustment indicated that the GARDENER and SCULPTOR metaphors

promoted the highest ratings of teacher responsibility, both significantly more than the COACH or

TOUR GUIDE metaphors, but not significantly different from one another. The COACH and TOUR

GUIDE metaphors did not significantly differ in terms of inferred teacher responsibility.

Metaphor also affected ratings of how much responsibility students had for their own learning,

$F(3, 393) = 11.91$, $p < .001$, $\eta^2 = .08$. Post-hoc comparisons indicated that the COACH and TOUR

GUIDE metaphors promoted the highest ratings of student responsibility, significantly more than

the GARDENER or SCULPTOR metaphors, but not different from one another. The SCULPTOR and

GARDENER metaphors were not significantly different from one another.
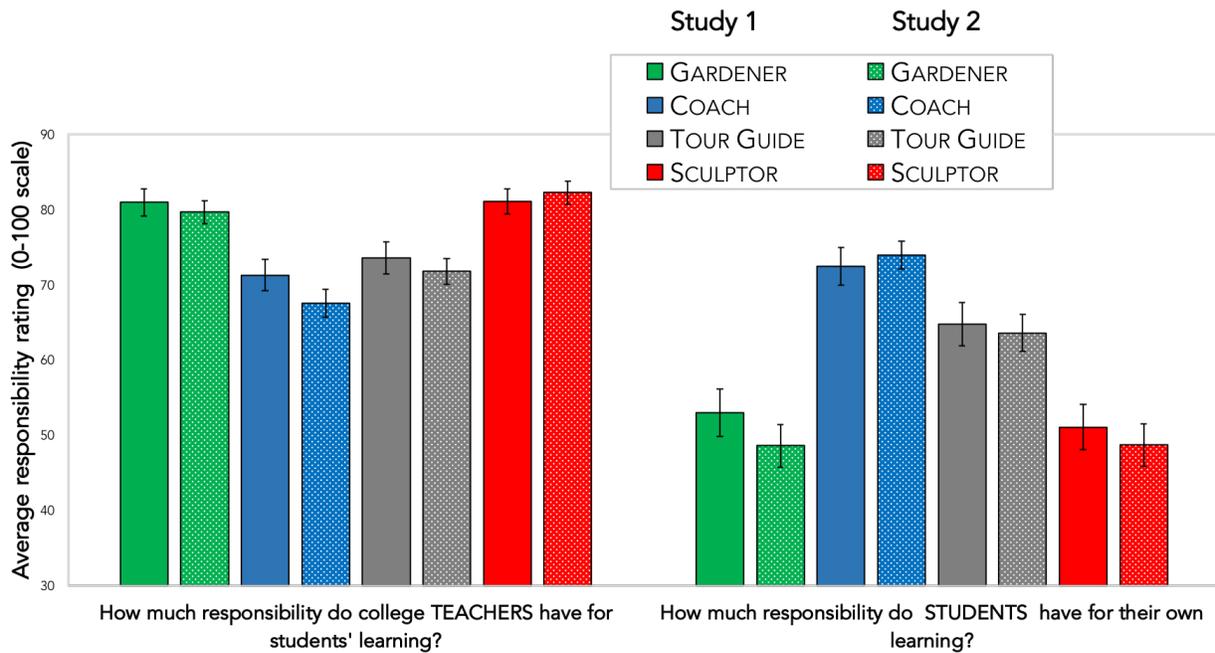

Figure 2. Mean ratings of teacher and student responsibility in Studies 1 and 2, as a function of the conceptual metaphor that participants considered. Study 1 results are shown in solid bars, Study 2 results in dotted bars. Error bars show +/- 1 SEM.

2.3.4. *Are different metaphors associated with different intuitions about teacher's power?*

Metaphor condition had a small but reliable effect on ratings of how much power college

teachers had to influence what students learn, $F(3, 391) = 5.08$, $p < .01$, $\eta^2 = .04$ (see Figure 3).

The *SCULPTOR* metaphor promoted the highest rating of teacher power over student learning, significantly more than the *COACH* metaphors, but not different from the *GARDENER* or *TOUR GUIDE* metaphors. The *GARDENER* metaphor was associated with higher ratings of teacher power over learning than was the *COACH* metaphor, but not significantly more than the *TOUR GUIDE* metaphor.

Metaphor had a more dramatic effect on ratings of how much power college teachers had to influence how students develop as people, $F(3, 391) = 20.42$, $p < .001$, $\eta^2 = .14$. Pairwise comparisons indicated that the *SCULPTOR* metaphor prompted the highest ratings of power to influence development, significantly more than either the *COACH* or *TOUR GUIDE* metaphor, but not significantly different from the *GARDENER* metaphor. The *GARDENER* and *COACH* metaphors were not significantly different from one another but both implied significantly more teacher power over development than did the *TOUR GUIDE* metaphor.
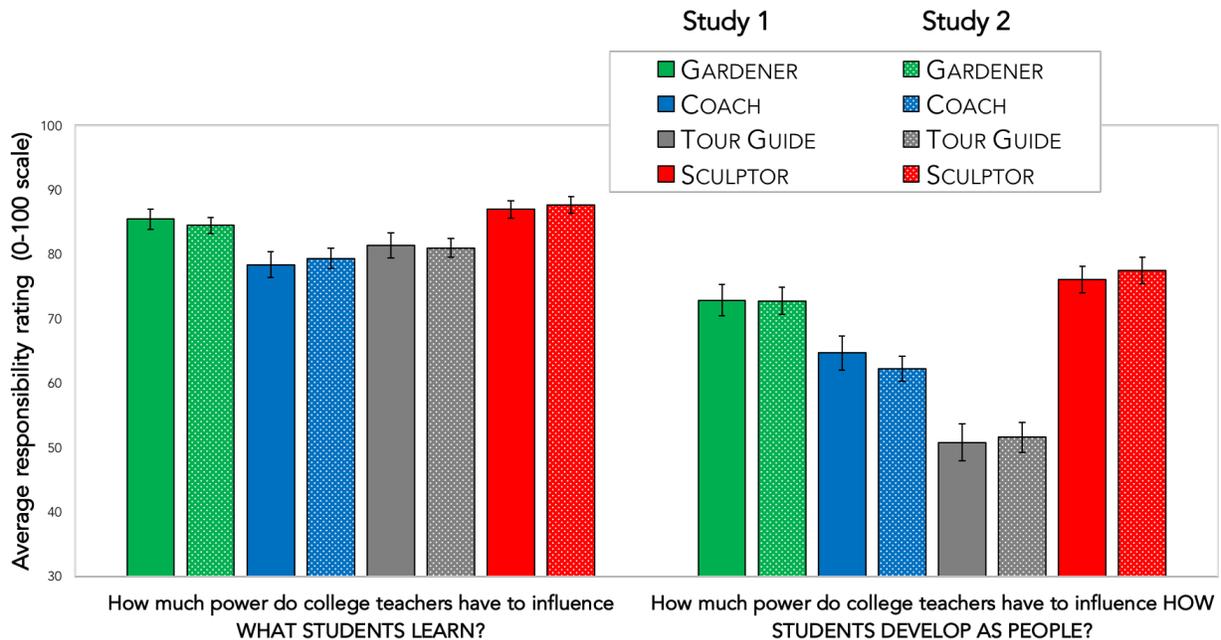

Figure 3. Mean ratings of teacher power to influence what students learn and how students develop as people, as a function of the conceptual metaphor that participants considered in Studies 1 and 2. Study 1 results are shown in solid bars, Study 2 results in dotted bars. Error bars show +/- 1 SEM.

2.3.5. *What was the relative popularity of the different metaphors?*

Of the four metaphors, the most popular was the GARDENER metaphor (33.4%), followed by the TOUR GUIDE (27.6%) and COACH metaphors (23.0%). The SCULPTOR metaphor was the least favored (13.9%). Only 2% of participants indicated preferring a metaphor other than these four.[3] See Table 4.

Table 4. Relationship between assigned metaphor condition and favorite metaphor in Studies 1 and 2. In study 1, participants' selection of favorite metaphor was not significantly influenced by the metaphor they were assigned. In study 1, assigned condition did influence favorite metaphor, such that participants were more likely to select the metaphor they were assigned as their favorite. This influence was strongest for the SCULPTOR metaphor, which was relatively unpopular except when participants were assigned that metaphor.

Favorite Metaphor	*GARDENER*		*TOUR GUIDE*		*COACH*		*SCULPTOR*		Total	
	Study 1	*Study 2*	*Study 1*	*Study 2*	*Study 1*	*Study 2*	*Study 1*	*Study 2*	**Study 1**	**Study 2**
GARDENER	31	44	33	30	32	38	36	38	**132**	**150**
TOUR GUIDE	29	42	22	49	28	37	30	29	**109**	**157**
COACH	20	40	28	41	20	48	15	37	**91**	**166**
SCULPTOR	12	7	15	10	8	9	20	31	**55**	**57**
Total	100	133	98	130	93	132	104	135	**387**	**530**

2.4. Discussion

In this study, participants were provided with one of four metaphors for teaching and rated the degree to which a series of 43 statements "fit" the metaphor. Using iterated exploratory factor analysis, we reduced the number of statements to just twenty, with four statements each loading uniquely on five latent factors—akin to teacher 'personality traits.' These five factors coherently varied in terms of endorsement for each of the four metaphors. This suggests that

different metaphors for teaching are associated with distinct entailments, and that different types of (metaphorical) teachers have different "profiles" or clusters of characteristics that can differ dramatically from one another.

Notably, the metaphors also differed in their implications for students' responsibility for their own learning. The COACH and TOUR GUIDE metaphors hold students more responsible for their own learning than do the GARDENER and SCULPTOR metaphors. The metaphors also differed in how much power the teacher has to influence a student's general development. The GARDENER metaphor entails the most power to influence students' development, and the TOUR GUIDE metaphor the least.

This study provides new insight into the underlying representational structure of different teaching metaphors, in addition to demonstrating a novel method for mapping metaphorical entailments. Because these findings are by their very nature exploratory, our goal for Study 2 was to replicate these patterns in a new sample and with a preregistered design and analysis.

3. Study 2

The goal of Study 2 was to replicate the essential findings of Study 1 using similar methods, but with our reduced set of 20 teacher characteristic statements. We hypothesized that the same five factor structure would emerge, and that these factors would vary across the four teaching metaphors in essentially the same way. We preregistered our study design and analysis plan, which can be found on the Open Science Framework at: https://osf.io/e6vba/.

3.1. Participants

We recruited 550 participants to complete the survey through Amazon's Mechanical Turk. Following our preregistered exclusion criteria, we excluded data from any duplicate IP addresses ($N = 0$), participants who were not between the ages of 18 and 30 ($N = 4$), had not attended college ($N = 1$), failed a memory check of which metaphor they had been assigned ($N = 10$), and/or provided duplicated (i.e., copy and paste) responses to all free response items as determined by the first author ($N = 1$). Although not preregistered, we excluded one additional participant who failed to provide ratings for any of the teacher attributes. See Table 1 for demographic data.

3.2. Materials & Procedure

The procedure was nearly identical to Study 1. Participants were randomly assigned to view one of the four teaching metaphors and were asked to think about the metaphor and write a sentence or two about how that metaphor could be true.[4] Participants then saw the reduced list of 20 statements describing college teachers derived from the iterated exploratory factor analysis conducted for Study 1. Participants rated how well each item fit the metaphor that they were given, and they were specifically instructed to rate how well each item agreed with the metaphor they received, not whether they personally believed each item was true. As in Study 1, participants answered four questions aimed at whether the different metaphors captured different intuitions about teacher responsibility and power. In contrast to Study 1, we added a memory check asking participants to identify which of the four metaphors they had been assigned, so that we could exclude data from participants who did not remember (and thus had not likely considered the assigned metaphor very carefully). Finally, participants viewed all four of the metaphors and selected their personal favorite and completed our demographics survey.

3.3. Results

3.3.1. *Replicating the factor structure of teacher characteristics.*

Our first step was to test whether the exploratory factor analysis performed in Study 1 would replicate in a new sample with a reduced set of 20 teacher attributes. Very simple structure and parallel analyses confirmed that five factors were a good fit for the data. We performed an initial principal axes factor analysis (oblimin rotation) in R, extracting five factors from the set of 20 attributes. The pattern of results was highly consistent with findings from Study 1, with the same 5 factors emerging.

The combined pattern matrix showing loadings for both Study 1 and 2 is shown in Table 2, correlations between factors in both studies are reported in Table 3, and the correlation matrix is available on the Open Science Framework. The factor loadings of each of the 20 teacher characteristics where highly similar to those found in Study 1, with a few small differences. The item "teachers are confident" loaded most strongly onto the engaging communicator factor in Study 1 (.61). Its loading on this same factor was noteworthy in the present study (.4), but it loaded slightly higher on the authority figure factor (.49). Additionally, the item "teachers stimulate students' thinking" loaded highly on the stimulator factor in Study 1 (.51) but failed to load highly (> .4) on this factor or onto any factor in Study 2.

3.3.2. *How do teacher characteristics vary across metaphors?*

Once a 5-factor solution was applied, we derived and saved the factor scores using the regression method in R. As in Study 1, metaphor condition (*GARDENER*, *COACH*, *TOUR GUIDE*, or *SCULPTOR*) had a significant effect on the set of factors, $F(15, 1530) = 18.9$, $p < .001$, $\eta^2 = .16$, with significant effects of metaphor on each factor individually: the *nurturer* ($F(3,512) = 18.83$, $p < .001$, $\eta^2 = .10$), *authority figure* ($F(3,512) = 14.87$, $p < .001$, $\eta^2 = .08$), *community-builder* (F

$(3,512) = 8.85$, $p < .001$, $\eta^2 = .05$), *engaging communicator* (F $(3,512) = 10.93$, $p < .001$, $\eta^2 = .06$), and *stimulator* factors (F $(3,512) = 2.71$, $p < .05$, $\eta^2 = .02$).

Figure 1 shows the pattern of differences across the metaphors for each factor in both Study 1 and 2. Visual examination of the figure shows a strikingly similar pattern of differences in teacher characteristics across metaphors in both studies.

3.3.3. *Are different metaphors associated with different intuitions about responsibility?*

Metaphor had a reliable effect on ratings of how much responsibility college teachers had for students' learning, $F(3, 532) = 17.0$, $p < .001$, $\eta^2 = .09$ (see Figure 2). Post-hoc tests using Bonferroni adjustment indicated that the GARDENER or SCULPTOR metaphors promoted the highest ratings of teacher responsibility, both significantly more than the COACH or TOUR GUIDE metaphors, but not significantly different from one another. The COACH and TOUR GUIDE metaphors did not significantly differ in terms of inferred teacher responsibility. Metaphor again affected ratings of how much responsibility students had for their own learning, $F(3, 532) = 23.9$, $p < .001$, $\eta^2 = .12$. Post-hoc comparisons indicated that the COACH metaphor promoted the highest ratings of student responsibility, significantly more than the TOUR GUIDE, GARDENER, or SCULPTOR metaphors. The TOUR GUIDE metaphor promoted higher ratings of student responsibility than did the GARDENER or SCULPTOR metaphors. The SCULPTOR and GARDENER metaphors were not significantly different from one another.

3.3.4. *Are different metaphors associated with different intuitions about teachers' power?*

Metaphor condition had a small but reliable effect on ratings of how much power college teachers had to influence what students learn, $F(3, 532) = 9.28$, $p < .001$, $\eta^2 = .04$. As shown in Figure 3, the SCULPTOR metaphor promoted the highest rating of teacher power over student

learning, significantly more than the *COACH* and *TOUR GUIDE* metaphors, but not different from the *GARDENER* metaphors. The *GARDENER* metaphor was associated with higher ratings of teacher power over learning than was the *COACH* metaphor, but not significantly more than the *TOUR GUIDE* metaphor. The *COACH* and *TOUR GUIDE* metaphor did not significantly differ.

Metaphor had a more dramatic effect on ratings of how much power college teachers had to influence how students develop as people, $F(3, 532) = 29.4$, $p < .001$, $\eta^2 = .14$. Pairwise comparisons indicated that the *SCULPTOR* metaphor prompted the highest ratings of power to influence development, significantly more than either the *COACH* or *TOUR GUIDE* metaphor, but not significantly different from the *GARDENER* metaphor. The *GARDENER* and *COACH* metaphors were not significantly different from one another but both implied significantly more teacher power over development than did the *TOUR GUIDE* metaphor.

3.3.5. *What was the relative popularity of the metaphors?*

In this sample, the *COACH* metaphor was most popular (31%), followed closely by the *TOUR GUIDE* (29%) and *GARDENER* metaphors (28%). The *SCULPTOR* metaphor was, once again, the least favored (11%). Only 1% of participants indicated preferring a metaphor other than these four.[5] See Table 4.

3.4. Discussion

The goal of Study 2 was to test whether the five-factor structure of teacher "personality" revealed in Study 1 would replicate using the reduced set of 20 teacher attributes, and whether these factors would differentiate the four teaching metaphors in the same way. Overall, the results indicated a largely successful replication. We observed the same set of five factors in the present study and the same teacher attributes loaded onto these factors, with only a few minor

deviations. These small differences from Study 1 may reflect less stability in two specific attributes ("teachers are confident" and "teachers stimulate students' thinking"). It is also possible that presenting participants with a smaller collection of teacher attributes (20 vs. 43) led them to perceive some attributes differently. Despite these small differences in the factor loadings, the factors themselves were differentially associated with the four metaphors in a very similar way as in Study 1. Similarly, differences in ratings of responsibility (student and teacher) and of teacher power across the four metaphors showed nearly identical patterns across the two studies.

4. General Discussion

People use metaphors to express their lay beliefs about everything from the nature of intelligence to how the economy works. In this paper, we focused on metaphors for college teaching and aimed to develop a systematic, principled method for mapping the entailments of metaphorical concepts, inspired in part by psychometric methods used to uncover the latent structure of psychological constructs like personality.

Across two studies, we examined four common, apt metaphors for the college teacher-student relationship: *TEACHER AS GARDENER*, *COACH*, *SCULPTOR*, and *TOUR GUIDE*, and found that these metaphors reflect systematically different intuitions about college teachers, captured by distinct, coherent clusters of teacher attributes. In the sections that follow, we discuss the implications of these findings for understanding lay beliefs about teaching, as well as the limitations and future directions for this work.

4.1. Metaphorical Conceptions of College Teachers

In Study 1, participants were provided with one of the four teaching metaphors and rated the extent to which a large set of teacher attributes conceptually cohered with the metaphor. We then used exploratory factor analysis to reduce the set of attributes to twenty, with four particular attributes each loading uniquely on a set of five meaningful latent factors. The five factors were akin to teacher personality traits, describing teachers as *nurturers*, *authority figures*, *community-builders*, *engaging communicators*, and *stimulators*. Further analyses revealed that these factors were differentially associated with each of our four teaching metaphors. The four metaphors were also associated with different implications for how much responsibility teachers and students hold for learning, and for how much power teachers have to influence learning and development more broadly. Study 2 comprised a preregistered replication of Study 1 focused only on the reduced set of twenty teacher attributes, and the results largely mirrored what we observed in Study 1, suggesting that our approach is relatively replicable.

Across both studies, the metaphor of college teachers as *GARDENERS* evoked a profile of teachers as nurturers who are relatively low on authority. The *GARDENER* metaphor also placed responsibility for student learning squarely on the teacher and suggested that teachers have great power to shape student learning and development. This popular metaphor differed dramatically in its conception of a college teacher compared to other popular metaphors. The teacher as *TOUR GUIDE* metaphor suggested teachers as engaging communicators who are relatively low nurturers, have just a little more responsibility for student learning than the students themselves, and have relatively little power to influence students. The *TEACHER AS COACH* metaphor elicited ratings of teachers as authority figures who build community, have slightly less responsibility for learning than students have, and moderate power to influence student learning and development. The

TEACHER AS SCULPTOR metaphor evoked teachers relatively high on authority and ability to stimulate student learning, and very high in power and responsibility for student learning.

4.2. Limitations and Future Directions

The present studies recruited online samples of current and former college students aged 30 and under. In future work, we plan to recruit more representative samples of current college students as well as college teachers and to explore whether preferred metaphors vary across demographic groups, academic disciplines, and educational settings (e.g., high school or elementary school). We will also examine whether preferred metaphors differ depending on whether participants are considering *ideal* versus *actual* college teaching. One limitation of the present work is that participants were asked to express their preferences for different metaphors, but whether the metaphor they selected described how college teachers *are* or *should* be was left open to interpretation.

Our findings also hint that different metaphors for teaching may be associated with differing expectations not only for how teachers are expected to behave but also for how much responsibility students, versus their teacher, hold for their learning (cf., Low, 2015; Seung, Park, & Jung, 2015; Wan, Low, & Li, 2011). Does the particular metaphor a student holds for the teacher-student relationship predict their attitudes and behaviors in the classroom? Based on our results, we would hypothesize that students who hold a *COACH* metaphor should expect teachers to be more demanding and assertive than students who hold a *GARDENER* metaphor, and thus may expect higher and stricter standards in the classroom. Similarly, because the *GARDENER* and *SCULPTOR* metaphors imply less responsibility on the part of the student, students who endorse these metaphors may hold a more passive view of the learning process and be less likely to

engage in active-learning strategies (e.g., self-quizzing) than students who endorse the *COACH* and *TOUR GUIDE* metaphors.

If metaphorical lay theories predict student and teacher behavior, then metaphorical language may offer a tool for shaping academic expectations and behaviors in ways that promote student success. Prior work suggests that even subtle differences in metaphorical framing can influence the solutions that participants generate to solve a complex problem (e.g., Thibodeau & Boroditsky, 2011). Can metaphorical language similarly influence how students and teachers conceptualize and approach the educational enterprise? For example, can instructors evoke the *COACH* metaphor (in syllabi or other communications) to encourage students to assume more responsibility for their own learning? Can college administrators evoke the *GARDENER* metaphor to encourage faculty to adopt a more "inclusive" approach to pedagogy? The Online Supplement includes a linguistic analysis of data from the present studies which yielded an initial library of metaphorical words associated with each of the four common teaching metaphors, and these may lay the foundation for experimental work to address these questions.

Ultimately, this research could lay the foundation for novel educational interventions based around metaphor framing. Some studies have found that metaphors can shape student mindsets (Blackwell et al., 2007) and attitudes (Landau et al., 2014), but this work has not specifically examined metaphors for teaching (see also Seung, Park, & Jung, 2015). As we described in the introduction, interventions aimed at changing intuitive lay theories of intelligence to improve student performance have generated inconsistent results (Sisk et al., 2018), but this may be in part to the fact that these interventions have neglected other important educational lay theories, like conceptions of teaching. As our current work offers a principled

way to understand the structure of people's intuitive beliefs about teaching, it provides an important first step in developing interventions that could address these limitations.

4.3. Conclusions

In sum, our findings demonstrate that (1) people endorse a variety of different teaching metaphors, (2) these metaphors bring to mind distinct, coherent clusters of teacher attributes, and (3) these metaphors are associated with different intuitions about teacher responsibility and power to shape student learning and development. Additionally, this work demonstrates a novel method for systematically mapping the entailments of metaphorical lay theories and sets the stage for educational interventions centered on shifting lay theories of teaching.

References

Alger, C. (2006). 'What went well, what didn't go so well': growth of reflection in pre-service teachers. *Reflective Practice*, *7*(3), 287–301.

Blackwell, L. S., Trzesniewski, K. H., & Dweck, C. S. (2007). Implicit theories of intelligence predict achievement across an adolescent transition: A longitudinal study and an intervention. *Child Development, 78*(1), 246-263.

Byrt, T., Bishop, J., & Carlin, J. B. (1993). Bias, prevalence and kappa. *Journal of clinical epidemiology*, 46(5), 423-429.

Chen, D. D. (2003). A Classification System for Metaphors about Teaching. *Journal of Physical Education, Recreation & Dance*, *74*(2), 24–31.

Chen, G., Faris, P., Hemmelgarn, B., Walker, R. L., & Quan, H. (2009). Measuring agreement of administrative data with chart data using prevalence unadjusted and adjusted kappa. *BMC medical research methodology*, 9, 5.

Davis, H. S., Watson, A. B., & Bakerson, M. (2015). Crowdchecking conceptual metaphors. In W. Wan & G. Low (Eds.), *Elicited Metaphor Analysis in Educational Discourse*, 139-165.

Dweck, C. S. (2008). *Mindset: The new psychology of success*. Random House Digital, Inc.

Flanagan, J. C. (1935). Factor analysis in the study of personality. Palo Alto, CA, US: Stanford Univ. Press.

Flusberg, S. J., & Hard, B. M. (2019). Metaphors we teach by: A method for mapping metaphorical lay theories. *Proceedings of the 41st annual meeting of the Cognitive Science Society, Montreal, Canada*

Gibbs, R. W. (1994). *The poetics of mind.* Cambridge University Press.

Gopnik, A., Meltzoff, A. N., & Kuhl, P. K. (1999). The scientist in the crib: Minds, brains, and how children learn. William Morrow & Co.

Grice, J. W. (2001). Computing and evaluating factor scores. Psychological Methods, 6(4), 430–450.

Gurney, B. F. (1995). Tugboats and tennis games: Preservice conceptions of teaching and learning revealed through metaphors. *Journal of Research in Science Teaching*, *32*(6), 569–583.

Keeley, J., Smith, D., & Buskist, W. (2006). The Teacher behaviors checklist: Factor analysis of its utility for evaluating teaching. *Teaching of Psych*, *33*(2), 84–91.

Keysar, B., & Bly, B. (1995). Intuitions of the transparency of idioms: Can one keep a secret by spilling the beans? *Journal of Memory and Language*, *34*, 89-109.

Kövecses, Z. (2005). *Metaphor in culture: Universality and variation*. Cambridge University Press.

Lakoff, G., & Johnson, M. (1980). *Metaphors we live by*. Chicago, IL: University of Chicago Press.

Landau, M. J., Oyserman, D., Keefer, L. A., & Smith, G. C. (2014). The college journey and academic engagement: How metaphor use enhances identity-based motivation. *Journal of Personality and Social Psychology, 106*(5), 679

Lee, K., & Ashton, M. C. (2007). Factor analysis in personality research. In R. W. Robins, C. Fraley, and R. F. Krueger (Eds.) *Handbook of research methods in personality psychology*, Guilford Publications.

Low, G. (2008). Metaphor and education. In Gibbs, R. W. (ed), T*he Cambridge handbook of metaphor and thought*, 212-231.

Low, G. (2015). A practical validation model for researching elicited metaphor. In W. Wan & G. Low (Eds.), *Elicited Metaphor Analysis in Educational Discourse*, 15–37.

Mahlios, M., Massengill-Shaw, D., & Barry, A. (2010). Making sense of teaching through metaphors: a review across three studies. Teachers and Teaching: theory and practice, 16(1), 49-71.

Murphy, G. L. (1996). On metaphoric representation. *Cognition*, *60*, 173-204.

Perkins, K., Toskos Dils, A., & Flusberg, S. J. (*under review*). The perceived threat of demographic shifts depends on how you think the economy works

Saban, A., Kocbeker, B. N., & Saban, A. (2007). Prospective teachers conceptions of teaching and learning revealed through metaphor analysis. *Learning and Instruction*, *17*(2), 123–139. doi: 10.1016/j.learninstruc.2007.01.003

Seferoğlu, G., Korkmazgil, S., & Ölçü, Z. (2009). Gaining insights into teachers' ways of thinking via metaphors. *Educational Studies*, *35*(3), 323–335. https://doi.org/10.1080/03055690802648135

Seung, E., Park, S., & Jung, J. (2015). Methodological approaches and strategies for elicited metaphor-based research. In W. Wan & G. Low (Eds.), *Elicited Metaphor Analysis in Educational Discourse*, 39–64.

Shaw, D. M., Barry, A., & Mahlios, M. (2008). Preservice teachers' metaphors of teaching in relation to literacy beliefs. *Teachers and Teaching: theory and practice*, *14*(1), 35-50.

Sim, J., & Wright, C. C. (2005). The kappa statistic in reliability studies: use, interpretation, and sample size requirements. Physical therapy, 85(3), 257-268.

Sisk, V. F., Burgoyne, A. P., Sun, J., Butler, J. L., & Macnamara, B. N. (2018). To what extent and under which circumstances are growth mindsets important to academic achievement? Two meta-analyses. *Psychological Science, 29*(4), 549- 571.

Thibodeau, P. H., & Boroditsky, L. (2011). Metaphors we think with: The role of metaphor in reasoning. *PloS one*, *6*(2), e16782.

Thibodeau, P. H., & Boroditsky, L. (2013). Natural language metaphors covertly influence reasoning. *PloS one*, *8*(1), e52961

Thibodeau, P. H., Crow, L., & Flusberg, S. J. (2017). The metaphor police: A case study of the role of metaphor in explanation. *Psychonomic Bulletin & Review*, *24*(5), 1375-1386.

Thibodeau, P. H., Matlock, T. & Flusberg, S. J. (2019). The role of metaphor in communication and thought. Language & Linguistics Compass, 13(5)

Tobin, K. (1990). Changing metaphors and beliefs: A master switch for teaching? *Theory Into Practice*, *29*(2), 122–127.

Wan, W., Low, G. D., & Li, M. (2011). From students' and teachers' perspectives: Metaphor analysis of beliefs about EFL teachers' roles. System, 39(3), 403-415

Yeager, D. S., & Dweck, C. S. (2012). Mindsets that promote resilience: When students believe that personal characteristics can be developed. *Educational Psychologist, 47(4),* 302-314

Yeager, D. S., Hanselman, P., Walton, G. M., Murray, J. S., Crosnoe, R., Muller, C., … Dweck, C. S. (2019). A national experiment reveals where a growth mindset improves achievement. *Nature*, *573*(7774), 364–369.

Footnotes

[1] Preliminary analyses of an initial sample were presented in the Proceedings of the 41st annual meeting of the Cognitive Science Society (Flusberg & Hard, 2019). That sample consisted of 199 participants ranging from 18-76 years of age ($M = 36$, $SD = 11.6$). Because conceptions of teaching have undoubtedly shifted over the past few decades, we wanted to restrict our age range to 18-30, which left only 82 participants from the initial sample. We recruited a second, larger sample of participants with the new age restriction, bringing our total sample size to close to 400 in the present Study 1.

[2] An exploratory factor analysis of the initial sample we collected was presented in the Proceedings of the 41st annual meeting of the Cognitive Science Society (Flusberg & Hard, 2019). There were two problems with this preliminary analysis that we have addressed in this study: First, in our earlier analysis we included all 43 teacher attribute items, some of which were highly skewed or loaded on one or all of the latent factors. Second, this earlier analysis produced a large and somewhat arbitrary number of latent factors, some of which did not differentiate between the metaphors. In consultation with a statistics expert, we addressed these issues in the present analysis, which includes an iterated exploratory factor analysis, dropping items based on skewed distributions and high loadings on multiple factors.

[3] A chi-square test of independence was conducted to determine if participants' preferred metaphor (excluding participants who selected "other) was influenced by condition. It was not, X^2 (9, $N = 387$) = 11.59, $p = .24$, suggesting that participants did not simply select a favorite

metaphor that matched what they had been presented with in earlier questions. This suggests that people may have stable preferred metaphors for thinking about the teacher-student relationship.

[4] The Online Supplement includes an analysis of the language that participants used in their responses to this question, aggregated across Study 1 and 2, with the goal of identifying common words and phrases associated with each metaphor.

[5] A chi-square test of independence indicated that participants' choice of favorite metaphor was influenced by the metaphor they had been assigned to consider previously received in earlier questions, X^2 (9, $N = 530$) = 35.8, $p < .001$. As shown in Table 4, preference for a given metaphor was higher when that metaphor had been assigned, an effect which was most dramatic for the SCULPTOR metaphor, which was far less favored in all conditions except the one in which it was assigned. While Study 1 suggested stability of preferred metaphors, the present study, with a larger sample, suggests that preferred metaphors are malleable and potentially influenced by experience reflecting on a given metaphor.

Supplemental Material

Note: This Supplemental Material provides additional information about a preliminary study conducted in preparation for the studies presented in the paper, as well as linguistic analyses of the free response data collected in the two studies.

1. Preliminary Study

In a preliminary study, we first identified and tested the relative popularity of six common metaphors for the college teacher-student relationship and identified the four most common. We derived these metaphors from qualitative findings of the metaphors that teachers use to describe their roles (e.g., de Guerrero & Villamil, 2002; Gurney, 1995; Oxford et al., 1998; Patchen & Crawford, 2011; Shaw et al., 2008). Each metaphor described both the teacher and the student and suggested a relationship between them (e.g. "*A teacher is like a sculptor and students are like clay*"). We assessed the relative aptness of these metaphors by asking participants to rate their agreement with each metaphor, select the one they liked best, and explain their choice. We also asked participants to justify their choices, and linguistic and thematic analyses revealed unique patterns of entailments associated with each metaphor. All measures used in the study and all relevant data are available through the Open Science Framework at: https://osf.io/e6vba/.

1.1. Participants

We crowdsourced 119 participants to complete an online survey through Amazon's Mechanical Turk. We required that participants be a current or former college student living in the U.S. or Canada, with an approval rating greater than 95% on at least 100 prior HITs. We

excluded data from 16 participants: two from duplicate IP addresses, 11 who included spam and/or duplicated (i.e., copy and paste) responses to all free response items, one whose responses indicated extremely weak proficiency in English, and two who reported having never attended college.

Of the 103 participants included in the final data set, 51% were male, 86% identified solely as White, 7% as Black, 3% as Hispanic/Latino, and 1% as Asian. The remaining participants identified as multiracial. Approximately 71% had graduated from college, 14% were currently enrolled students, and 16% had attended college at one point but were not currently enrolled. About 40% of those who had attended college attended less than 4 years ago and 60% had been out of college for more than 4 years. Mean age was 35 ($SD = 11$) with a range of 20-74.

1.2. Materials & Procedure

We designed a survey using Qualtrics survey software which presented participants with six possible metaphors for the teacher-student relationship (see Table S1). We derived these metaphors from qualitative findings of the metaphors that teachers use to describe their roles (e.g., de Guerrero & Villamil, 2002; Gurney, 1995; Oxford et al., 1998; Patchen & Crawford, 2011; Shaw et al., 2008). Each metaphor described both the teacher and the student and suggested a relationship between them (e.g. "*A teacher is like a sculptor and students are like clay*"). All of these metaphors originated in previous literature with the exception of the *app store* metaphor, which was a contemporary adaptation of a "*provider of tools*" metaphor (de Guerrero & Villamil, 2002) and "*business-oriented*" metaphor (Chen 2003).

Participants viewed all six metaphors in a random order and each was described as a metaphor for college teaching. Participants first rated their agreement with the metaphor on a 6-

point scale (strongly disagree to strongly agree, with no neutral midpoint) and then freely explained their response by typing into a text box. After considering all six metaphors, participants considered the entire collection of metaphors and selected their favorite. They explained why they preferred this metaphor and, specifically, how it conformed to their experiences and views of college teaching. Finally, participants answered a series of basic demographic questions.

All measures used in the study and all relevant data are available through the Open Science Framework at: https://osf.io/e6vba/.

1.3. Results

1.3.1. *What are the most popular metaphors?*

As shown in Table S1, participants leaned toward agreement ($M > 3.5$) for all the metaphors except the *app store* metaphor, but most strongly endorsed the *gardener* metaphor, followed closely by the *coach*, *tour guide*, and *sculptor* metaphors. Participants' selection of their preferred metaphors showed a similar pattern of popularity. Of the six metaphors, the most popular favorite was the *gardener* metaphor (34.0%), followed by the *tour guide* (19.4%) and *coach* metaphors (18.4%). The *sculptor* (9.7%), *ship captain* (8.7%), and *app store* metaphors (7.8%) were less favored.

1.3.2. *How did participants extend metaphors?*

Our next goal was to examine whether and how participants extended the metaphors that they were given. First, we examined the degree to which participants used metaphorical language in justifying their agreement with each metaphor. The second author examined the justifications

that participants offered for their agreement with each of the six metaphors and rated them for metaphoricity on a 1 to 5 scale with 5 representing the most metaphorical (e.g., *"The teacher molds students and shows them new shapes. Students are raw materials to be shaped and guided."*) and 1 representing literal language or language that was nearly identical to the prompt (e.g.,*"Teacher is like a sculptor and student is like a clay to teach to succeed in our life."*). Some responses were deemed uncodable due to their incomprehensibility and were excluded from further analyses (e.g., *"NA"*). A total of 598 responses were codable, and 95 participants had codable responses across all six metaphors. A second, independent judge also rated the metaphoricity of the descriptions. Overall interrater reliability was computed using a bivariate correlation and was fairly high ($r = .633$, $p < .001$). The average metaphoricity rating was 2.87 out of 5 ($SD = 1.17$).

We performed a repeated-measures ANOVA on the metaphoricity ratings across all six metaphors and found significant variation, $F(5,465) = 9.17$, $p < .001$, $\eta^2 = .09$. As shown in Table S1, metaphoricity ratings were highest for the *gardener* metaphor, followed by *ship ship captain*, *sculptor*, *coach*, *tour guide*, and finally the *app store* metaphors. For all but the *sculptor* metaphor, agreement and metaphoricity were positively correlated, as shown in Table S1. Participants who agreed more with a given metaphor more likely to extend the metaphor in their justification.

Our second step was to identify the common and unique words associated with each metaphor. We used Python's Natural Language Toolkit (nltk) library to perform frequency counts and n-gram analyses for responses to each metaphor (after appropriate text preprocessing and excluding any words that were also present in the prompt). Table S2 shows the most common words that were found in a minimum of 1% of descriptions for each metaphor, ordered

from most to least frequent. Table S2 also showed the most common unique words for each metaphor that were found in a minimum of .1% of descriptions. Common unique words were words found only for that metaphor and not for the others.

As Table S2 shows, many of the common unique words were metaphorical in nature, drawn from the metaphorical source to talk about teaching. For example:

The *gardener* metaphor prompted descriptions of teachers who "sow information" and plant "seeds of knowledge" to "cultivate students" who can "flourish and grow" "through to fruition."

The *tour guide* metaphor led to descriptions of teachers who "take you around" into "the unknown" and let students' "minds wander" along "the path of learning" in the "world of knowledge."

The *coach* metaphor elicited descriptions of teachers facilitating a classroom where students "work together" as "part of the team" to "develop skills and knowledge" including "a plan of attack" in pursuit of "a common goal."

The *sculptor* metaphor prompted descriptions of teaching as "hands-on," and described students as being "raw material," an "unfinished block" or even "blank slates" that can become sculpted products to "reveal what is hidden."

The *ship captain* metaphor yielded descriptions of a teacher who "gives orders" with the "sailor being guided" while "teachers must discipline" in a "complicated hierarchy" all while "driving their ship."

The *app store* metaphor encouraged descriptions of as how students "pick and choose" their topics of study in a unique "learning environment" that functions in an educational "operating system" in which they "store all the content."

These descriptions suggest that participants were actively using the metaphor to reason about the qualities of a teacher associated with a given metaphor (Thibodeau et al., 2017).

1.3.3. *How were metaphors thematically different?*

The previous analysis revealed common words and phrases associated with each metaphor for the teacher-student relationship, but what were the broader conceptual associations for each? More specifically, what attributes of instructors and students come to mind for each metaphor as participants are justifying how much they agree with it? To find out, the third author performed a thematic analysis of participants' justifications for their agreement with each metaphor. Using an inductive approach, we identified preliminary themes describing implications of the metaphor for instructor or student characteristics and then focused on the nine themes that were present in at least 10% of the responses for a given metaphor. A second, independent coder assessed the free responses for the same themes. Interrater reliability using Cohen's Kappa ranged from 0.180 to 0.719 ($M = 0.544$). Because many of the themes were relatively low in frequency for a given metaphor, we also assessed interrater reliability using PABAK (Prevalence-Adjusted-Bias-Adjusted Kappa) (Byrt et al., 1993; Chen, Faris, Hemmelgarn, Walker, & Quan, 2009). According to this measure, reliability ranged from 0.656 to 0.900 ($M = 0.814$).

Participants broadly described each metaphor's implications for instructor influence, student effort, the self-directed nature of student learning, instructor care, and students' personal growth. Notably, participants sometimes described how a metaphor implied the presence of a certain quality (e.g. student effort), and sometimes its *absence.* These were classified as different themes. For example, participants sometimes described how a metaphor was associated with a

high degree of students' effort (e.g., "The coach expects his athletes to work hard and perform well when it counts."). In other cases, participants described how a metaphor was associated with a *lack* of student effort (e.g., "This makes the students sound completely passive, which I don't like. Education is a result of the teacher AND the student cooperating. Students have choices and different interests and aren't just molded by the teachers.").

More detailed descriptions for each theme along with definitions, examples, prevalence, inclusion and exclusion criteria, and measures of interrater reliability for each theme are available through the Open Science Framework at: https://osf.io/e6vba/. Next, we examined the relative frequency of each theme across the six teaching metaphors, as well as the correlation between the presence of that theme and participants' agreement rating for that metaphor. Table S3 displays the results. Most themes correlated positively with metaphor agreement. For example, themes of instructor care correlated with higher ratings of agreement for the *gardener* metaphor (e.g., "*The professor has plenty of resources needed to help students. If students are like the plants than they can live off of the professor to get what they need and to grow*.") However, some themes negatively correlated with agreement. In all cases, this was because participants who rated lower agreement with a metaphor were more likely to mention the theme as their justification. The theme of instructors' lack of influence, for example, correlated with lower agreement ratings for *tour guide* metaphor. In other words, participants disagreed with this metaphor because it implied that teachers had little influence. For example, one participant wrote, "No. Teachers actually teach people the things that they need to learn, not just show things and shuffle along."

1.4. Discussion

The findings from this preliminary study suggest that people with college experience endorse a variety of metaphors for the teacher-student relationship, but the most apt are the *gardener*, *coach*, *tour guide*, and *sculptor* metaphors. The results also provided some initial evidence that these metaphors are associated with distinct conceptual entailments, as the themes that emerged in participants' response justifications revealed a variety of coherent relationships between specific metaphors and specific qualities of teachers. For example, participants who agreed with the *tour guide*, *sculptor*, and *coach* metaphors were likely to note that instructors influence their students, while those who agreed with the *gardener* metaphor were more likely to note that teachers care about their students and that students undergo personal growth. That participants spontaneously extended the metaphors in articulating their personal preferences provides some evidence that they were *thinking* about the nature of college teaching in terms of the metaphorical lay theory, using the metaphors as cognitive tools to scaffold their reasoning.

Although a useful starting point for understanding metaphorical lay theories for teaching, this initial analysis was more suggestive than systematic: the themes we derived were created post-hoc during analysis and provided a very coarse differentiation between the metaphors. As an analogy, this method is somewhat akin to creating personality trait profiles by asking people to describe themselves and then deriving traits that distinguish between people based on themes that emerge from their responses. Although a suggestive and possibly fruitful first step, this method is subject to researcher bias and unlikely to yield a valid or reliable taxonomy of traits.

2. How did participants extend metaphors in Studies 1 and 2?

Mirroring the analyses in our preliminary study, we sought to examine the language used to extend each of the assigned metaphors in Studies 1 and 2. Participants were asked to write a

sentence or two about how a college teacher could be like the metaphor that was assigned. We applied natural language processing to these data, aggregated across the two studies, with the goal of identifying common metaphorical words and phrases associated with each metaphor. We used Python's Natural Language Toolkit (nltk) library to perform another exhaustive frequency count and n-gram analyses for responses to each metaphor (after appropriate text preprocessing and excluding any words that also featured in the prompt). Table S4 shows the most common words and unique words found in a minimum of 1% and .1% of descriptions for each metaphor, respectively. As in the preliminary study, many of the common unique words were metaphorical in nature: participants used words from the source domain (e.g., water, seed, tend, cultivate) to describe the target domain of teaching.

Table S1. For each metaphor in the preliminary study, the mean ratings of agreement (6-point scale), metaphoricity (5-point scale), and their correlation.

Metaphor	Mean Agreement Rating (SD)	Mean Metaphoricity (SD)	Correlation between Metaphoricity and Metaphor Endorsement
GARDENER: *"A teacher is like a gardener and students are like plants"*	4.34 (1.28)	3.39 (1.17)	.293**
COACH: *"A teacher is like a coach and students are like athletes"*	4.24 (1.24)	2.82 (1.15)	.203*
TOUR GUIDE: *"A teacher is like a tour guide and students are like tourists"*	4.12 (1.32)	2.69 (1.18)	.240*
SCULPTOR: "A teacher is like a sculptor and students are like clay"	4.07 (1.24)	2.96 (1.12)	-.057^{ns}
SHIP CAPTAIN: "A teacher is like the CAPTAIN of a ship and students are like sailors"	3.69 (1.39)	3.01 (1.02)	.330***
APP STORE: "A teacher is like an app store and students are like smartphone users"	2.81 (1.51)	2.51 (1.19)	.384***

^{ns} not significant **;***$p < .05$; **$p < .01$; ***$p < .001$

Table S2. Most common words and unique words used to justify agreement with each metaphor in the preliminary study. Words are listed in descending order of frequency (most to least). Words with tied scores are presented alphabetically.

Metaphor	Most Common Words (Descending ≥ 1%)	Most Common Unique Words (Descending ≥ .1%)
GARDENER: Total Words = 1284	grow(5.84), help(2.34), make(1.56), good(1.40), metaphor(1.40), need(1.40), knowledge(1.32), agree(1.01), one(1.01), tend(1.01)	garden(.78), seed(.62), cultivate(.47), feed(.47), flower(.31), health(.23), watch(.23), weed(.23), cannabis(.16), careful(.16), continue(.16), distinction(.16), flourish(.16), food(.16), limit(.16), nourishment(.16), nutrient(.16), sculptor(.16), sow(.16), wither(.16)
COACH: Total Words = 1225	teach(2.12), guide(1.55), work(1.55), metaphor(1.47), team(1.47), think(1.47), help(1.06)	motivate(.57), athletic(.41), performance(.33), player(.33), physical(.24), toward(.24), win(.24), wisdom(.24), blame(.16), compete(.16), exercise(.16), freely(.16), habit(.16), receive(.16), require(.16), urge(.16)
TOUR GUIDE: Total Words = 1255	show(3.27), thing(2.15), learn(1.91), think(1.35), metaphor(1.27), know(1.12), knowledge(1.12)	city(.32), explore(.32), visit(.24), casual(.16), deep(.16), detail(.16), expose(.16), inform(.16), lose(.16), sight(.16), surface(.16), tangible(.16), temporary(.16), village(.16), visitor(.16),
SCULPTOR: Total Words = 1397	mold(3.58), sculpt(1.50), think(1.43), teach(1.36), professor(1.29), metaphor(1.15), shape(1.15), learn(1.07), knowledge(1.00), make(1.00)	sculpt(1.50), product(.29), turn(.29), create(.21), finished(.21), lump(.21), molding(.21), personality(.21), anyone(.14), artist(.14), blank(.14), emphasis(.14), entirely(.14), malleable(.14), piece(.14), raw(.14), ready(.14), rough(.14), slate(.14), slightly(.14), stone(.14), unformed(.14), wheel(.14), wood(.14)
SHIP CAPTAIN: Total Words = 1290	work(1.71), follow(1.63), guide(1.63), learn(1.63), lead(1.55), metaphor(1.47), leader(1.40), classroom(1.24), teach(1.24), class(1.16), one(1.16), order(1.09), think(1.01)	crew(.62), boss(.31), move(.31), figure(.23), follower(.23), sail(.23), else(.16), main(.16), nobody(.16), obey(.16), overseer(.16), port(.16), represent(.16),
APP STORE: Total Words = 1216	make(2.14), information(2.06), metaphor(2.06), use(2.06), choose(1.64), think(1.64), knowledge(1.56), really(1.40), thing(1.40), get(1.32), learn(1.23), sense(1.23), one(1.15)	access(.66), phone(.49), datum(.33), service(.33), search(.25), cell(.16), connection(.16), contain(.16), device(.16), download(.16), meaningful(.16), operate(.16), opposite(.16), technological(.16), whenever(.16)

Table S3. Most common themes used to justifications for ratings of agreement with each metaphor in the preliminary study, as well as the prevalence and correlation of the theme with agreement. Themes that were present in at least 10% of responses are shown in order of most to least common. Negative correlations with agreement indicated that the theme was more likely to be found in justifications for low agreement with a given metaphor.

Metaphor	Most Common Themes in Justifications for Ratings of Agreement	Prevalence (% of Responses) (> 10% of responses; ordered most to least common)	Correlation with Agreement
SCULPTOR	Instructors' influence	71%	.472***
	Students' personal growth	17%	.289**
	Students' lack of effort	14%	-.415***
	Students' self-directed learning	11%	-.225*
GARDENER	Instructors' care	70%	.498**
	Students' personal growth	62%	.436**
	Instructors' influence	44%	.128^{ns}
	Student's lack of self-directed learning	12%	-.434**
COACH	Instructors' influence	73%	.388**
	Students' effort	34%	.176^{ns}
	Students' personal growth	26%	.314**
	Instructors' care	15%	.098^{t}
	Students' self-directed learning	14%	-.078^{n}
APP STORE	Students' self-directed learning	31%	.324**
	Students' effort	26%	.270**
	Instructors' lack of Influence	23%	-.260**
	Instructor's influence	18%	.533**
	Instructors' lack of care	11%	-.181^{t}
	Instructor's influence	59%	.531**

TOUR GUIDE	Students' effort	20%	.282**
	Students' self-directed learning	20%	.264**
	Instructors' lack of Influence	18%	-.492**
	Students' lack of effort	16%	-.511**
SHIP CAPTAIN	Instructor's influence	66%	.614**
	Students' effort	33%	$-.147^{ns}$
	Students' lack of self-directed learning	23%	-.328**

^{ns} not significant **;***$p < .05$; **$p < .01$; ***$p < .001$

Table S4. Most common words and unique words used to describe each assigned metaphor in Studies 1 and 2. Words are listed in descending order of frequency (most to least). Words with tied scores are presented alphabetically.

Metaphor	Most Common Words (Descending ≥ 1%)	Most Common Unique Words (Descending % ≥ .1%)
GARDENER: *"A teacher is like a* ***gardener*** *and students are like plants"* Total Words = 2945	grow(7.61), help(3.74), knowledge(3.02), water(2.00), college(1.87), care(1.77), give(1.36), need(1.32), learn(1.26), seed(1.12), take(1.09)	water(2.00), seed(1.12), nurture(.98), tend(.92), cultivate(.61), flourish(.58), nourish(.37), nutrient(.31), thrive(.31), bloom(.24), nourishment(.24), mature(.20), blossom(.17), flower(.17), food(.17), garden(.17), healthy(.17), survive(.17), die(.14), prune(.14), raise(.14), sunlight(.14), daily(.10), extra(.10), nature(.10), soil(.10)
COACH: *"A teacher is like a* ***coach*** *and students are like athletes"* Total Words = 2862	college(2.41), learn(1.71), teach(1.64), train(1.50), guide(1.43), good(1.29), help(1.29), game(1.08), well(1.05), work(1.01)	train(1.5), game(1.08), sport(.87), play(.66), team(.66), win(.45), training(.35), performance(.31), player(.21), victory(.17), athletic(.14), couch(.14), exam(.14), fundamental(.14), rule(.14), body(.10), common(.10), compete(.10), equip(.10), feedback(.10), instead(.10), limit(.10), physically(.10), strategy(.10), yeah(.10)
TOUR GUIDE: *"A teacher is like a* ***tour guide*** *and students are like tourists"* Total Words = 3109	learn(2.38), show(2.35), subject(2.06), information(1.93), college(1.83), knowledge(1.61), new(1.51), know(1.45), teach(1.25), thing(1.25), topic(1.22), help(1.09), way(1.06)	explore(.58), familiar(.35), question(.35), unfamiliar(.32), visit(.29), location(.19), territory(.19), city(.16), highlight(.16), interested(.16), overview(.13), walk(.13)
SCULPTOR: *"A teacher is like a* ***sculptor*** *and students are like clay"* Total Words = 3196	mold(3.41), shape(2.16), mind(2.10), college(2.00), teach(1.75), way(1.72), knowledge(1.47), think(1.38), make(1.35), help(1.28), learn(1.22), sculpt(1.22)	sculpt(1.22), sculpture(.31), malleable(.22), art(.19), piece(.19), masterpiece(.16), power(.16), slate(.16), belief(.13), carve(.13), style(.13)