2018

Evaluating Teacher Beliefs and Behaviors: Identifying Teacher "Types"

Madison Holzman

Pete Swerdzewski

Annika Jessen

Evaluating Teacher Beliefs and Behaviors: Identifying Teacher "Types"

Madison A. Holzman
Peter Swerdzewski
Annika Jessen

Curriculum Associates

Abstract

Data are increasingly used in the modern K–12 classroom. Educators indicate that using data to inform instruction is a necessary component of effective teaching (DQC, 2018). *i-Ready* is an educational product designed to assist teachers with collecting and using data to inform curricular changes. Teachers' beliefs regarding *i-Ready*, as well as how they use *i-Ready* data, may be factors in important student learning outcomes. In this initial study, a latent profile analysis (LPA) was conducted to evaluate "types" of teachers based on their beliefs and behaviors regarding *i-Ready*. Four classes of teachers were identified: believers, users, neutralists, and compliers. Teachers were found to vary more in their beliefs than their behaviors, and teachers who were neutral in their beliefs about *i-Ready* were found to review and discuss data less than all other teachers. Future studies will be geared toward collecting validity evidence to support the four classes.

Introduction

Role of Data in the Modern Classroom

Data are abundant in today's education system. To meet accountability mandates, states must provide information about student learning. With the movement toward standards-based education over the past decade, states must provide information regarding the extent to which students meet the state-adopted standards of learning. Because federal money and other resources are linked to student performance on end-of-year assessments, educators face pressure for students to meet proficiency standards.

To prepare students for the summative end-of-year assessments, educators often use formative assessments to gauge students' abilities throughout the academic year and guide classroom instruction (DQC, 2018). Specifically, teachers may use formative assessment data to identify struggling students and create an alternative learning pathway for struggling or below-grade level students. Alternatively, formative assessment data may be used to identify students who are performing above grade level and create an alternative learning pathway to challenge those students. Formative assessment data are used by educators to create custom learning experiences for individual students as well as guide whole class curriculum.

Though data can be useful in the modern classroom, their benefits are only realized by the review and intentional use of data to inform instruction. Unfortunately, educators indicate that they do not use data to their full potential. Limited training is an often-described barrier to using formative assessment data to inform instruction. Additionally, teachers say that even if they received training and resources for how to use data, they do not often have the time to thoroughly review data and plan intentional curricular changes (DQC, 2018; Means, Padilla, DeBarger, & Bakia, 2009). Thus, though teachers view data use as an important aspect of being an effective educator (DQC, 2018), using data is challenging and not always intuitive. Such findings are problematic, as reviewing and using formative data is often a key aspect of curriculum implementation. Without the time or knowledge of how to use the data, educators may be missing an important curricular factor related to student learning.

To assist educators, education companies provide products and tools designed for curriculum planning and provide training for reviewing and using data to inform instruction. One such product, Curriculum Associates' *i-Ready* program, is the focus of this paper.

i-Ready

i-Ready is a suite of instructional materials and a formative assessment tool that provides a multifaceted experience for educators and students. In addition to a core curriculum of math or English language arts (ELA), students may use *i-Ready Instruction* for approximately 45 minutes per subject per week to supplement their core curriculum. *i-Ready* was developed based on the College and Career Readiness Standards, as well as other state standards, thereby providing students with practice related to state learning standards.

The *i-Ready Diagnostic* is a computer-adaptive, formative assessment taken by students at three times during the academic year: fall, winter, and spring. From the Diagnostic, teachers receive information about students' on-grade level status (e.g., three or more grade levels below, two grade levels below, one grade level below, on grade level, one grade level above, two or more grade levels above) and students' domains of strength and weakness. Importantly, teachers receive information about their entire class as well as individual student scores and on-grade level information. Both types of data can inform large- or small-group instruction as well as individual instruction. Based on the Diagnostic results, students are routed to a series of *i-Ready Instruction* lessons, of which students spend approximately 45 minutes per week working through. *i-Ready Instruction* lessons are designed to supplement students' learning above and beyond their core math or English language arts curriculum. From *i-Ready Instruction*, teachers receive information about students' time on task, lessons completed and passed, and domains of strength and weakness. As in the Diagnostic, teachers receive information about their entire class as well as individual students' progress with *i-Ready Instruction* lessons.

To assist with interpreting data reports and provide guidance on effectively using data in the classroom, districts implementing *i-Ready* receive professional development from their district administrators and/or Professional Development Specialists at Curriculum Associates. To evaluate teachers' perceptions and uses of *i-Ready*, Curriculum Associates surveys educators with a biannual Educator Survey.

Educator Survey

The Educator Survey is administered twice a year to teachers, school administrators, and district administrators. Educators are randomly selected to receive the survey, and educators receive a small monetary incentive to complete it. Though the survey is administered biannually, educators only receive the survey once. Thus, all responses to the Educator Survey are from unique educators, and data are not longitudinal.

The Educator Survey is comprised of Likert-response, rank-order, select all that apply, and open-ended questions designed to gauge educators' satisfaction with *i-Ready*, beliefs regarding *i-Ready* and *i-Ready* data, and behaviors regarding the use of *i-Ready* data. Questions are related to *i-Ready* as a suite, as well as specifically related to the *i-Ready Diagnostic* or *i-Ready Instruction.* Educator Survey responses have proven useful to understand how *i-Ready* beliefs and behaviors may vary by length of time using *i-Ready*, types of professional development educators receive, and the grades with which educators are using *i-Ready*. This information can inform resource allocation, particularly regarding the types of professional development to provide in the future, or whether to target teachers of particular grades for additional support and training.

Though the Educator Survey yields useful information for understanding how teachers view and use *i-Ready,* the Educator Survey yields a lot of information. There are many beliefs- and behavior-related questions to sift through, and the amount of data is sometimes challenging to pare down to make concise statements regarding teachers' beliefs and behaviors, and how these beliefs and behaviors may relate. As such, Curriculum Associates explored a latent profile analysis method to identify profiles for teacher beliefs and behaviors.

Identifying Response Profiles to Help Classify Teacher Types

Creating profiles of teacher types is a useful technique for understanding teacher beliefs and behaviors. Specifically, because the creation of profiles is used to group teachers based on similar response patterns, researchers can identify if there are distinct differences between teachers' beliefs and behaviors. For example, the creation of profiles helps researchers identify whether teachers with high positive beliefs also report using data more often than other teachers, or if there is a subpopulation of teachers who have high positive beliefs, yet do not use the data. Profile creation allows for more nuanced interpretations of teachers' beliefs and behaviors than descriptive or simple regression methods allow.

Perhaps most importantly, identifying profiles of teachers provides an opportunity to explore how teachers' beliefs and behavior patterns may relate to important outcomes such as student growth or student proficiency. By evaluating such relationships, researchers can direct resources to guide teachers toward beliefs and/or behaviors that may most likely yield high student growth and greatest student proficiency. For example, suppose it is found that students working with teachers who have negative beliefs about *i-Ready* and *i-Ready* data demonstrate the least positive growth compared to students who work with teachers who have positive beliefs. In this case, it may be a beneficial use of resources to work with teachers who have negative beliefs and actively attempt to improve their beliefs. On the contrary, if students were found to demonstrate the same growth regardless of teacher beliefs, then researchers could dedicate resources to other initiatives to improve students' learning, as improving teachers' beliefs may not yield improvements in student growth. Identifying how teacher profiles relate to student learning is not only beneficial for Curriculum Associates to better understand how our products relate to student learning, but also for education as a whole as we attempt to learn more about how students learn and how to improve educational experiences for students.

There is interest at Curriculum Associates in better understanding the profiles or "types" of teachers using *i-Ready*. Eventually, Curriculum Associates staff hope to evaluate whether teachers with certain beliefs and/or behaviors are associated with more student growth or differential satisfaction. The current project was a preliminary analysis designed to address the following questions:

1) Do teachers systematically differ in their beliefs and behaviors regarding *i-Ready*? Specifically, when considering teacher beliefs and behaviors, are there distinct "types" of teachers?
2) If there are distinct "types" of teachers, how do these "types" of teachers differ in the beliefs and/or behaviors regarding *i-Ready*?

Method

To address these questions, a latent profile analysis (LPA) was conducted on teacher responses from the Educator Survey. An LPA is a type of mixture modeling in which analyzed data are continuous in nature (Masyn, 2013). LPAs are conducted with the assumption that an overall distribution of responses is comprised of underlying classes, with each class having a distribution that contributes to the shape of the overall response distribution. In instances of multi-modal and/or non-normal response distributions, LPAs may be particularly useful, as the non-normal shape of the overall response distribution could be due to underlying classes with their own response distributions.

Teacher responses from the December 2017 administration of the Educator Survey were used for the analysis. Analyzed data were from 857 teachers, and all teachers had complete data on the questions included in the LPA. Responses to five behaviors questions and twelve beliefs questions were used in the analysis (see Appendix A for the questions involved in the analysis). All questions were five-point Likert-type questions. The behaviors questions' response options ranged from either *"Never"* to *"Weekly"* or "*Never*" to *"Daily*." The beliefs questions' response options ranged from "*Strongly Disagree*" to "*Strongly Agree*." All data were treated as continuous in the analysis.

One- through five-class models[1] were estimated using *Mplus* version 8.1 (Muthén & Muthén, 2017). For all models, covariances between questions were constrained to zero, in effect assuming no correlation between questions. The means and variances for each question were freely estimated. Note that it is not necessarily realistic to assume no correlation between questions; however, because this was a preliminary analysis and means were the main information of interest, we were comfortable with constraining covariances for this analysis.

To evaluate convergence, the condition number and log-likelihood values were examined.[2] To evaluate model fit, the Bayes Information Criteria (BIC; Schwarz, 1978), Sample-

[1]When using maximum likelihood estimation, it is imperative to find a global maximum to ensure that the solution for which the data are most likely has been identified. As such, mixture models are typically estimated using several thousand start values to ensure that a global, rather than local, maximum has been identified. All models were estimated twice, once with 1,000 start values, and again with 2,000 start values to ensure the lowest log-likelihood values were identified and replicated. All results are from the analysis with 2,000 start values.

size adjusted Bayes Information Criteria (SS-BIC; Sclove, 1987), Bayes Factor, and Lo-Mendell-Rubin Likelihood Ratio Test (LMR-LRT; Lo, Mendell, & Rubin, 2001) were evaluated. For both the BIC and SS-BIC, smaller values suggest better model fit. The Bayes Factor compares the BIC between two models, and values greater than one suggest a simpler model (e.g., a c-1 class model) fits the data better than a more complex (e.g., c class) model (Wasserman, 2000). The LMR-LRT provides a test of the null hypothesis that a more complex model (i.e., c class model) does not fit the data better than a less complex model (i.e., a c-1 class model). A significant LMR-LRT suggests the more complex model fits the data better than the simpler model (Lo et al., 2001). Though all model-fit statistics were considered, emphasis was placed on the SS-BIC and LMR-LRT when determining model fit (Tofighi & Enders, 2007). In addition to model fit indices, classification accuracy and proportion of sample in each class were considered. Classification accuracy for individual classes was evaluated, as well as entropy, an average estimate of classification accuracy across classes.

Results

Results from the five-class model are unavailable due to convergence issues. Fit statistics for the one- through four-class models are provided in Table 1. All fit indices suggested the four-class model provided the best fit to the data. Moreover, classification accuracy was high (>93%) for the four-class model. Sample proportion for the four-class model was adequate, though the first class is rather small at only 39 teachers (see Table 2). A four-class model was championed, but results should be interpreted cautiously as the fourth class may be a result of over-extraction of classes.

[2]Log-likelihood values were evaluated to ensure that the lowest log-likelihood value was replicated. The condition number was monitored to ensure empirical identifiability. Values less than 10^{-6} may indicate empirical underidentification and unstable solutions (Muthén & Muthén, 2017). In this study, all condition numbers were greater than 10^{-3}.

Figure 1 demonstrates the response patterns for each class of the championed four-class model. Questions to the left of the vertical dashed line represent behaviors questions. Questions to the right of the vertical dashed line represent beliefs questions (for the wording of each question, see Appendix A). The classes are perfectly ordered across all beliefs questions; however, classes one and two flip ordering for the behaviors questions. Moreover, there is more variability between classes for beliefs than behaviors. That is, teachers differ more in their beliefs than their behaviors. Even those teachers in class one who tend to disagree with beliefs questions review data nearly monthly. Those teachers in class two who tend to respond neutrally review data the least out of all classes of teachers.

Discussion

Because LPA is a largely exploratory technique, researchers are indebted to collect validity evidence to support the classes extracted via LPA. Because no validity evidence was gathered for this analysis, results should be interpreted cautiously. With this limitation in mind, four classes of teachers were identified via this LPA. These classes are tentatively named the compliers (class one), neutralists (class two), users (class three), and believers (class four). The compliers are characterized as having the lowest beliefs about *i-Ready* compared to all other teachers, yet are still reviewing online instruction data and discussing data with their principals, colleagues, and students nearly monthly. Thus, even though these teachers do not have positive beliefs regarding *i-Ready*, they are still using data fairly often, potentially from a compliance aspect. The neutralists are characterized as having neutral beliefs about *i-Ready*, and they review Online Instruction data and have discussions about data with their principals approximately once per quarter. Neutralists review data with their colleagues and students nearly monthly, similarly

[2]Log-likelihood values were evaluated to ensure that the lowest log-likelihood value was replicated. The condition number was monitored to ensure empirical identifiability. Values less than 10^{-6} may indicate empirical underidentification and unstable solutions (Muthén & Muthén, 2017). In this study, all condition numbers were greater than 10^{-3}.

to the compliers. The users are characterized as having positive beliefs about *i-Ready,* and they review Online Instruction data and have conversations about *i-Ready* data with principals and colleagues monthly, similar to the compliers. However, the users review data with students more often than the compliers. The believers are characterized as having the most positive beliefs about *i-Ready* compared to all other teachers, and they review *i-Ready* data the most.

An advantage of LPA is the identification of clear patterns of responses. From LPA results, it is clear there are four distinct “types” of teachers. Moreover, given that the classes are not perfectly ordered due to the switch in rank-order of compliers and neutralists in terms of their behaviors, there is evidence that classes do represent true distinctions in teacher “type.” That is, teachers do not only differ in their degree of beliefs and behaviors (i.e. perfect rank-order across questions), but compliers and neutralists differ qualitatively when considering their beliefs and behaviors. These patterns would be challenging to detect via traditional descriptive or simple regression techniques. With LPA results, Curriculum Associates staff can identify how teachers’ beliefs and behaviors relate to one another and how to respond to teachers of various classes.

How to respond to various classes may vary depending on the end goal. For example, upon review of results for the compliers or neutralists, the initial reaction may be to work with compliers and neutralists to improve their beliefs regarding *i-Ready*. However, whether this action is a beneficial use of resources may depend on the outcome of interest. If the end goal is student learning, it will be interesting to evaluate how students grow when working with compliers compared to users or believers. If students working with compliers grow as much as students working with users or believers, it may not be necessary to address the low beliefs of compliers. However, if the end goal is *i-Ready* satisfaction, knowledge of neutralists and compliers provides a clear avenue for action, as additional resources could be dedicated to these teachers, specifically to better understand and address their neutral or negative beliefs. Knowledge of classes provides an opportunity for Curriculum Associates staff to relate the classes to various outcomes to better understand resource allocation.

Though knowledge of the classes is useful, what is more useful is *why* these patterns are emerging. Identifying classes is an important first step and opens the door for follow-up inquiry. Qualitative inquiry may be particularly useful when further understanding these classes. For example, though we tentatively named class one the "compliers," are the teachers in this class really using the data from a compliance perspective? Or is there some other reason why these teachers are using data fairly often, despite negative beliefs regarding *i-Ready*? Qualitative inquiry can help us also understand what has made teachers in class four have such positive beliefs regarding *i-Ready.* Are these teachers primarily receiving certain types of professional development? Do their districts have characteristics that assist in a smooth implementation of *i-Ready*? Knowing more about the believers and users may provide information that can be used when working with neutralists or compliers.

Next Steps

As mentioned, this was a preliminary study with limitations. Several follow-up studies will be conducted, primarily to identify a model to allow for the ordinal treatment of behaviors responses, as well as incorporate variables for validity evidence into the model. One validity variable may be the professional development teachers received. Curriculum Associates has studied how professional development relates to teachers' beliefs and behaviors. From previous studies, we know that teachers who receive some professional development have more positive beliefs and use the data more than teachers who do not receive professional development. A follow-up analysis may include professional development type to evaluate whether teacher classes align with our previous knowledge related to professional development.

Additionally, recall that covariances between questions were constrained to be zero. Because responses to questions are likely correlated with one another, a series of models with varying degrees of estimated covariances should be tested to determine the extent to which responses should be able to covary.

Finally, all teachers in this study had complete data. That is, data were listwise deleted if teachers were missing data on any of the included beliefs or behaviors questions. Listwise deletion assumes data are missing completely at random, a stringent assumption that is not likely met in practice (Enders, 2010). When the missing completely at random assumption is not met, results may be biased in unpredictable ways. As such, in follow-up studies, teachers who have missing data should be included in the analysis, and an appropriate estimator, such as full-information maximum likelihood, should be used.

Conclusion

In this study, we evaluated latent classes of teachers via a latent profile analysis (LPA). Teachers were found to vary more in their beliefs than their behaviors. Interestingly, even those teachers with the lowest beliefs still reviewed data nearly monthly. Before using these classes to make decisions about resources or evaluate how the classes may relate to student learning, validity evidence should be gathered to support the classes. In short, LPAs are a useful technique for educational research, as results can help us better understand our teachers, students, and how we can improve student learning.

References

DQC (2018). What parents and educators think about education data. Washington DC: Data Quality Campaign.

Enders, C. K. (2010). Applied missing data analysis. New York, NY: Guilford Press.

Lo, Y., Mendell, N. R., & Rubin, D. B. (2001). Testing the number of components in a normal mixture. *Biometrika, 88,* 767–778.

Masyn, K. E. (2013). Latent class analysis and finite mixture modeling. In T. D. Little (Ed.), *Oxford library of psychology*. *The Oxford handbook of quantitative methods: Statistical analysis* (pp. 551–611). New York, NY: Oxford University Press.

Means, B., Padilla, C., DeBarger, A., & Bakia, M. (2009). Implementing data-informed decision making in schools: Teacher access, supports, and use. Report prepared for the U.S. Department of Education, Office of Planning, Evaluation, and Policy Development. Menlo Park, CA: SRI International.

Muthén, L. K. & Muthén, B. O. (2017). *Mplus user's guide* (8th edition). Los Angeles, CA: Muthén & Muthén.

Schwartz, G. (1978). Estimating the dimension of a model. *The Annals of Statistics, 6,* 461–464.

Sclove, L. S. (1987). Application of model selection criteria to some problems in multivariate analysis. *Psychometrika, 52,* 333–343.

Tofighi, D., & Enders, C. K. (2007). Identifying the correct number of classes in growth mixture models. In G. R. Hancock (Ed.), *Mixture Models in Latent Variable Research* (pp. 317–341). Information Age: Greenwich, CT.

Wasserman, L. (2000). Bayesian model selection and model averaging. *Journal of Mathematical Psychology, 44,* 92–107.

Table 1.
Fit statistics for one- through five-class models

	BIC	SSA-BIC	LMR-LRT	BF	Entropy
One-class	39113.36	39005.38	--	--	--
Two-class	35505.86	35340.72	<0.001	<0.001	0.922
Three-class	33494.40	33272.10	<0.001	<0.001	0.953
Four-class	32102.91	31823.44	<0.001	<0.001	0.943
Five-class	--	--	--	--	--

Note. BIC = Bayes Information Criteria; SSA-BIC = Sample-size Adjusted Bayes Information Criteria; LMR-LRT = Lo-Mendell-Rubin Likelihood Ratio Test; BF = Bayes Factor

Table 2.
Proportion of sample for four-class model

	N	Proportion
Class One	39	0.046
Class Two	192	0.224
Class Three	476	0.555
Class Four	150	0.175

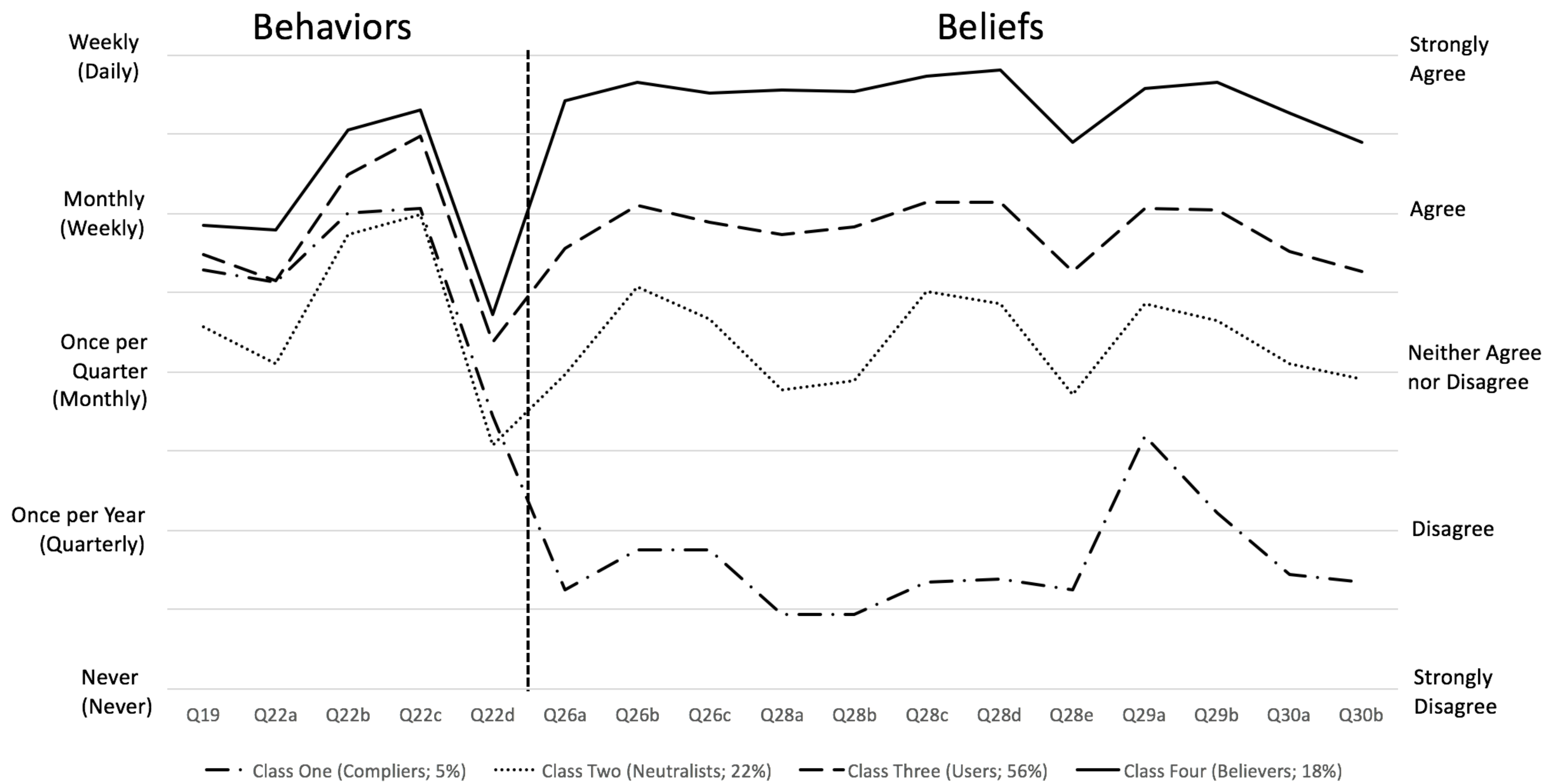

Figure 1. Response patterns for four-class model. Response options for behaviors questions are in left-hand axis, whereas response options for beliefs questions are in right-hand axis. Response options for behaviors questions in parentheses represent the response options for Q19; all other behaviors questions use the response options not in parentheses.

Appendix A
Educator Survey Questions Included in the LPA

Behaviors	
Question	**Question Wording**
Q19	When do you review *i-Ready* data about students' progress on Online Instruction?
Q22a	How often do you have conversations about *i-Ready* data with your principal and/or coach?
Q22b	How often do you have conversations about *i-Ready* data with your colleagues?
Q22c	How often do you have conversations about *i-Ready* data with your students?
Q22d	How often do you have conversations about *i-Ready* data with your students' families?

Note. The response options for Q19 were: Daily, Weekly, Monthly, Quarterly, Never. The response options for all other behaviors questions were: Weekly, Monthly, Once per Quarter, Once per Year, Never.

Beliefs	
Question	**Question Wording**
Q26a	*i-Ready* is a tool that helps me be a better teacher.
Q26b	*i-Ready* helps me differentiate instruction to meet the needs of all students.
Q26c	*i-Ready* helps me address the on-grade level instructional needs of my classroom.
Q28a	*i-Ready* assessments are trustworthy.
Q28b	*i-Ready* assessments provide reliable data.
Q28c	*i-Ready* data helps me know how my students are doing.
Q28d	*i-Ready* data helps me understand my students' needs.
Q28e	*i-Ready* data helps me know how my students will do on state assessments.
Q29a	*i-Ready* Online Instruction is rigorous and standards-aligned.
Q29b	Using *i-Ready* Instruction leads to student growth.
Q30a	*i-Ready* helps students understand their academic progress, encouraging them to take more ownership over their learning.
Q30b	*i-Ready* helps families understand their students' academic progress.

Note. The response options for all beliefs questions were: Strongly Agree, Agree, Neither Agree nor Disagree, Disagree, Strongly Disagree.

Holzman, M. A., Swerdzewski, P. S., & Jessen, A. (2018, December). *Evaluating Teacher Beliefs and Behaviors: Identifying Teacher "Types"* (Curriculum Associates Research Report No. RR 2018-37). North Billerica, MA: © Curriculum Associates, LLC. Prepared 12/12/2018.