

## A survey to measure perceived agency of teachers

Michael M. Hull\* 

University of Vienna, Austria  
michael.malvern.hull@univie.ac.at

Haruko Uematsu

Tokyo Gakugei University, Japan  
uematsu@u-gakugei.ac.jp

### Abstract

Dissemination of reformed curriculum requires teachers to feel that they have the freedom to actually implement the curriculum. Whereas greater learning gains are found with active learning, this generally requires more time than rote lecturing. National standards which call for wide content coverage make such approaches prohibitive. This and other pressures may lead teachers to feel that they have little control over what and how they teach. Here, we assess a survey we designed to measure a “teacher’s perceived agency”, which we define as “a feeling of being in control over what is taught and of how it is taught.”

### Keywords

agency, Rasch model, teacher attitude, teacher training

♦Received 20 January 2020 ♦Revised 5 May 2020 ♦Accepted 17 June 2020

### Introduction

Physics education research has established that traditional physics teaching, where the teacher lectures while the students quietly take notes, is ineffective from a variety of perspectives. Generally speaking, students learn painfully little content knowledge (e.g. Hake, 1998), their attitudes about the nature of physics knowledge and learning tend to get worse (e.g., Redish, Saul, & Steinberg, 1998), and they come to find the subject to not be interesting. Research has led to the development of curricular materials and teaching strategies that have been demonstrated to be more effective, and this has affected the training that pre-service teachers (PSTs), like those at the University of Vienna (UV) and at Tokyo Gakugei University (TGU) receive. In addition to learning content knowledge, such as Maxwell’s equations and Newton’s laws of motion, PSTs learn pedagogical knowledge, such as classroom management and assessment strategies, and pedagogical content knowledge, which includes description of student naive ideas in physics and effective strategies for helping students learn particular topics. Certainly, if we wish students to be taught with the most effective means available, it is necessary for their teachers to first learn the material and how to teach it themselves. We argue, however, that knowledge is not sufficient. Even if a teacher understands reformed curriculum sufficiently to be able to apply that knowledge in the classroom setting, what ultimately matters is whether or not the teacher actually does apply it.

### State of the literature

- Teacher’s perceived agency has influence on what teachers do in the classroom.
- Instilling in pre-service teachers a sense of agency is an instructional goal for some teacher educators.
- Teachers’ perceived agency has been studied qualitatively via interviews and in-class observations. This allows for deep understanding of individuals, but generalization across and between populations of teachers (for example, before and after a professional development workshop) would benefit from a survey. To date, none exists.

### Contribution of this paper to the literature

- We created an original survey to measure a teacher’s perceived agency. Five survey validation interviews and a panel consisting of 10 experts suggest that 31 items on the survey are valid. Responses from 3 pre-service teachers who took the survey twice indicate that the instrument is also reliable.
- Rasch analysis results from 144 pre-service teachers’ further support the argument that 27 of those items are valid and reliable in distinguishing respondents based upon their level of perceived agency.
- Although our survey is specific to physics, the questions themselves can readily be modified for use in other contexts.

What are the factors that affect whether research-based curricula and pedagogical approaches are disseminated or not? Perhaps the most immediately-available explanation lies with the pedagogical knowledge or beliefs of the teachers themselves. Many physics teachers resist reformed approaches to teaching physics because they assume that the traditional instruction that was effective for them will be effective also for their students. However, even teachers who are convinced of the value of research-based instruction and are trained in it as PSTs might fail to implement the curriculum in the classroom if their school or colleagues do not support their attempts. Biesta et al. write that many teachers are faced in the school with a “mishmash of competing and vague ideas – personalization, choice, learning, subjects, etc.”, “are regularly left confused about their role”, and hence tend to think more about short-term obligations and less about the long-term purposes of education (Biesta, Priestley, & Robinson, 2015). We can thus imagine a teacher who thinks “well, I know that it would be better for these students to learn this topic using the technique I learned in teacher-training, but doing that would take up more time, and we have a tight schedule.” In many countries, the class time necessary to cover the wide breadth of topics put forth by national standards makes the use of research-based curriculum prohibitive, as interactive engagement typically requires more time (Ball & Cohen, 1996; Biesta et al., 2015; Edwards & D’arcy, 2004; Wei & Chen, 2019). Teachers are influenced not only by what administrators say, but also by parents (Ball & Cohen, 1996) and their perceptions of their students (Ball & Cohen, 1996; Biesta et al., 2015). For example, teachers may believe that students prefer traditional teaching, and this might prevent them from using interactive engagement

methods (Biesta et al., 2015). An additional reason a teacher may hesitate to use reformed curriculum is consideration of the status quo (Biesta et al., 2015). In Japan, for example, it is not uncommon for a school to resist change, as a result of inertia. Consequently, high school physics teachers graduating from TGU have reported feeling pressure to teach in the traditional style used by other teachers at their school instead of with the research-based curriculum they learned as PSTs. If a teacher feels that colleagues or even the principal is hostile towards the reformed instruction in which the teacher was trained, then this will naturally hinder the dissemination of the new material.

Strictly speaking, however, it is not the control that the teacher does or does not have that affects his or her actions, but rather the teacher's perception of that control. In the most dramatic case, even if the school is, in fact, completely indifferent to what the teacher does in the classroom, if the teacher fears consequences (which, in this hypothetical situation, do not actually exist) for not teaching as the prior teacher did, then the reformed curriculum will still not be disseminated. We conceptualize a teacher with a strong sense of agency as one who, amongst a backdrop of various external influences, nevertheless perceives control of what is taught and of how it is taught, and we define "perceived agency" in this way. That is, we define "perceived teacher agency" to be "a feeling of being in control over what is taught and of how it is taught." In this paper, we will report on the validation process of a survey we designed to gain insight into this trait. Although the survey is specific to physics instruction, we suspect that it can readily be modified for use in other contexts. For example, Question 1 (Q.1) reads "I will consider carefully what physics textbook to use in my classroom." This could readily be changed to instead reference "history textbook", etc.

The Perceived Agency Survey has the teacher (either a PST or an in-service teacher) consider various situations to probe whether the teacher feels in control of his or her own teaching, or whether the teacher feels controlled by a range of factors, including national standards, colleagues, and even the textbook and other curricular materials used in the lesson themselves. Sawyer discusses some curricula, such as Success for All, which "provide word-for-word scripts that teachers are strongly encouraged to follow" in order to "teacher-proof" instruction (Sawyer, 2004). Considering as well the exhausting course load that high school teachers in many countries face, there are sentiments that it would be best to have teaching physics require minimal innovation by the teacher. With this model, the teacher would faithfully utilize curriculum developed by education researchers without a need for taking the time to interpret or understand the rationale behind the curriculum. Many education researchers, however, argue that responsive teaching is crucial (e.g., Debarger et al., 2017; Harlow, 2009, 2010; Robertson, 2018; Sawyer, 2004; Schrittmesser, 2013). Sawyer, for example, writes "...the best teachers apply immense creativity and profound content knowledge to their jobs, both in advance preparation and from moment to moment while in the classroom" (Sawyer, 2004). Such a perspective has implications for instruction of PSTs. Researchers have argued, for example, that to help PSTs bridge the gap

between theory and practice (where creativity is required), that PSTs should be encouraged to draw upon their own experiences and ideas and to connect it with what they are learning (e.g., Engle & Faux, 2006). One potential benefit of our survey is that it can provide a quantitative, and hence comparative, measure of the effectiveness of such instruction.

### *Theoretical background*

Although the concept of "agency" has been extensively discussed, particularly in sociological literature, relatively little has been written about teacher agency (Biesta et al., 2015). We do not mean "agency" in the sense of being an extension of an organization, like a secret agent. Rather, we mean something very similar to what Enghag and Niedderer describe as "ownership", which "refers to the importance and need... to actually participate by discussions, choice, responsibility, and decision taking" (Enghag & Niedderer, 2008). Milner-Bolotin (Milner-Bolotin, 2001) describes "ownership of the learning process" for learners in project-based-instruction group projects in terms of three overlapping realms, with full overlap corresponding to most ownership: 1) does the learner find personal value in the project (e.g., finds the knowledge to be useful and/or connected to prior knowledge?); 2) does the learner feel in control to make decisions and be proactive?; and 3) does the learner feel responsible for the learning process and results of the project? We mean something very similar by "perceived agency" of a teacher, but we avoid the word "ownership" as, when used in the context of teaching rather than learning, it conjures up images of strict disciplinarian and/or traditional lecture-based forms of teaching that an agentive teacher may choose to avoid.

Our usage of the phrase "perceived agency" is consistent with the definition of "agency" as "the capacity to initiate purposeful action that implies will, autonomy, freedom, and choice" (Lipponen & Kumpulainen, 2011). In the case of agency in the context of learning, Davydov et al. defined "agent" as "a person who knows how to teach herself, determines the limits of her own knowledge (ignorance), and finds for herself the means to expand the limits of the known, of the accessible" (Davydov, Slobodchikov, & Tsukerman, 2003). We might paraphrase this as something like "a person who is in control over what is learned and how it is learned." We similarly define "perceived (teacher or PST) agency" to be "a feeling of being in control over what is taught and of how it is taught." A teacher or PST with a weak sense of agency, then, feels controlled in this regard, either by the education system (Ball & Cohen, 1996; Biesta et al., 2015; Edwards & D'arcy, 2004; Wei & Chen, 2019), colleagues at the school (Biesta et al., 2015), or by other factors including student (Ball & Cohen, 1996; Biesta et al., 2015) and parent (Ball & Cohen, 1996) expectations.

In situations where teachers feel like they are being controlled, it is easy to place blame on overly-demanding national curricula or old-fashioned schools. Researchers who study agency in a general sense, however, argue that one mischaracterizes the situation in thinking of agents as being directly controlled by the societal structures in which they act. In writing about the agency of children in pursuing science as a career path, for example, Archer et al. described two children,

Vanessa and Poppy, who were both encouraged by the structure put in place by their parents to pursue science, but nevertheless exercised “individual agency in developing and asserting alternative choices”(Archer et al., 2012). Researchers describe an interplay between the agents and the structure, with the views of the agents being affected, but not completely controlled, by the structure, and the agents then responding in ways that either support or challenge the structure (Archer et al., 2012; Olitsky, 2006; Sewell, 1992; Smith, 2013). In much the same way, it should be expected that two teachers in the same school could have different levels of perceived agency and hence might have different propensities towards enacting reformed curricula. Although we consider research comparing the structures themselves, for example, between countries, to be of vital importance (e.g., Bao et al., 2009; Hull, 2013; Tosa, 2011), we see the question of what teachers do within those structures, that is, their agency, as also being a relevant subject of study. In choosing a survey format for our research, we recognize that we cannot speak to what decisions teachers actually make in the classroom—we gain only limited insight into the “capacity” of a teacher or PST “to initiate purposeful action that implies will, autonomy, freedom, and choice” (p. 813 of Lipponen & Kumpulainen, 2011). Rather, our instrument enables one to make claims about how teachers perceive this capacity.

### Perceived Agency Survey Design

Although a search online revealed no surveys to measure teacher agency specifically, two surveys were found which measure agency in other contexts, and these served as a basis for the survey we created. The first survey, the Ownership Measurement Questionnaire, was created by Milner-Bolotin to measure the feelings and beliefs of non-science majors who worked on a group project for a physical science course (Milner-Bolotin, 2001). The second survey, Perceived Choice and Awareness of Self Scale, includes questions which measure whether or not one perceives a sense of choice behind his or her actions (“Perceived Choice and Awareness of Self Scale (PCASS),” 2019). Both surveys utilize a five-point Likert scale, as does the survey we created. As an example, one of the prompts from the Ownership Measurement Questionnaire was “I feel responsible for the in-depth exploration of the project-topic” (p. 172 of Milner-Bolotin, 2001), which is similar to Question 42 (Q.42) of our Perceived Agency Survey: “I feel responsible for making my students think deeply.” In cases where an item on one survey was very similar to an item on the other survey, we condensed the two items into a single item. We added questions as needed to capture the range of factors discussed in literature that may hinder teachers in expressing their agency (see above). We wished to have an equal number of positive as negative statements, so this led to the creation of additional statements. For example, Q.42 is the pair of Q. 22: “It is not my job to make students think deeply—that is their decision.” Finally, to keep the survey from being too long, we removed items that we predicted to not be useful and/or to be redundant with other items already on the survey. For example, we first modified the prompt “I have a sense of ownership of the group-project I am working on” from the Milner-Bolotin survey to be “I have a sense of ownership of my teaching” and, conversely “I do not have a sense of ownership of my teaching”. We then considered that students would likely either be confused by these

prompts or that it would be redundant with the responses to the other prompts, and so we removed both prompts. The second author translated this survey into Japanese for administration at TGU.

### Survey Validity and Reliability

The first author administered this first draft of the Perceived Agency Survey (in English) to PSTs at the start of the 2018 summer semester at UV (Hull and Uematsu, 2020). After students had completed the survey (20 minutes were allowed), the first author had the PSTs discuss four of the prompts, first in pairs and then in whole-class discussion. We chose these four prompts in particular because we anticipated that they would be the most relevant to the PSTs and lead to the most generative discussion. Prior to this discussion, the PSTs signed consent forms allowing the discussion to be videotaped. We present bits of transcript below to triangulate with responses PSTs selected on the survey as an indicator of the validity of the survey even at this developmental stage. This discussion was also formative for us, in that it indicated some areas of the survey that were being poorly understood, indicating need for improvement. For example, some students interpreted the word “curriculum” in the original prompt “Curriculum is developed by experts, and it should be used without messing it up” to mean “national syllabus”. Since the intent of the question was instead to ask about material which might be used to satisfy the national syllabus (such as a PER-based worksheet or a textbook), this question was replaced with five new questions, Q.1, Q.6, Q.13, Q.20, Q.33 (see **Table 1** below).

Overall, it seemed from this discussion that the UV PSTs in this class felt a rather strong sense of agency, a finding that was consistent with the responses to the survey. In response to the prompt, “If the principal of my school tells me to teach in a certain way, I will do my best to teach that way, even if I don’t really want to,” Adler (all PST names are pseudonyms) replied

*I put “disagree”, because I feel, if I don’t think that this is the best way to teach my students, how should the students get the value of it? If I myself disagree with the subject I’m teaching? For me it’s hard to teach something if I’m not believing in it, not really thinking “this is making sense, this is a good way of teaching it”, for me it is hard to provide that. I think it is like lying to yourself. But I only put “disagree” there and not “strongly disagree”, because, at the end of the day, the principal is still our boss; we are not working independently. So I am under them, and, of course, they are the boss, they need to inform what I do in my classroom.*

When given the prompt “Research-based textbooks are developed by experts, and they should be used without messing it up,” two students responded

*Billie: Most textbooks I studied in the last few years had extreme faults in them, so it’s not that they know the answer to everything, even if they are good. It’s not always possible to find a perfect textbook, so you have to mess with it a little bit, because every class is different, so you can’t do exactly what they would have you do.*

*Dresden: I think it is seldom a good idea to use just one source for those kinds of things, so probably you should take multiple books and maybe compare or whatever.*

Finally, in response to the prompt, “My students will have taken many classes before taking my class, and they will have an idea of how a class ‘should go’. I need to teach in that style too, otherwise it will be too strange for my students,” Billie responded

*I mean, it’s really difficult, just saying that you have to do it the same as the other teacher because I’m not another teacher, I’m me, and I have to stay me even if I have a class that is not used to how I teach. If I tried to be someone else, the class wouldn’t be authentic. So I don’t think it would be good for the class.*

Taken collectively, the utterances of these PSTs indicate that they are thinking that they will rely upon their own views about what is the best way to teach and that they will be proactive in deciding which curriculum to use and how to implement it. As discussed in (Hull and Uematsu 2020), this is consistent with the responses selected by the PSTs on the survey. Encouraged by this pilot study, we then further developed the survey into a 44-item instrument for additional assessment of both validity and reliability.

The 44-item survey was translated into German in the fall 2018 semester by a native German speaker, who then used it to conduct a succession of survey validation interviews which led to a finalized survey (Hull et al., 2019). These interviews not only finessed the German translation of the survey, but also led to 13 questions being changed on the English and Japanese versions as well. These questions are indicated with italics in **Table 1**.

**Table 1.** The Perceived Agency Survey (31 questions remaining after the expert panel).

Item #	Item
1	I will consider carefully what physics textbook to use in my classroom.
2	If the principal of my school tells me to teach in a certain way, I will do my best to teach that way, even if I don’t really want to.
3	If my physics students do not understand what they are learning, I will take more time with the material, even if that means that some planned topics are not taught in class.
4	I prefer curriculum that tell the teacher exactly what to do, so that I don’t risk making the wrong decision.
6	<i>I will just use whatever physics textbook the teacher before me used. If it was good enough for him/her, then it is good enough for me.</i>
9	It might be the case that at my school where I am teaching, a more experienced teacher will not want me to use research-based pedagogy but to instead stick to traditional ways of teaching. Nevertheless, I will keep trying to introduce curriculum that I think will be the most effective.
13	<i>Once I choose a physics textbook, I will just use it, at most, as a guide. I will not hesitate to skip sections or point out to students which parts I think are poorly-worded, confusing, or wrong.</i>

**Table 1 (Continued)**

14	Teaching is just a job so I can get a paycheck – there is no benefit to me beyond that.
15	Outdated equipment at my school is not an excuse for a poor lesson. I will just have to rely more on creativity!
16	It doesn’t really matter whether I do my part in helping students learn or not—they will meet plenty of other teachers.
17	I feel that I have control over what I teach and how I teach it.
18	What my students learn in my class will have little benefit for them in other courses and/or in everyday life.
19	I will provide quality education to my students, even if I need to spend more time preparing for class than my colleagues do.
20	Once I choose a physics textbook for my classroom, I will follow it carefully.
21	<i>I think I have influence over the progress of my students.</i>
23	I will teach in the way I think is best, regardless of what my principal or other teachers might think.
27	<i>Curricular resources are, at most, a guide for teachers to use or modify creatively, as the situation requires.</i>
28	My students will have taken many classes before taking my class, and they will have an idea of how a class “should go”. I need to teach in that style too, otherwise it will be too strange for my students.
29	Parents should not tell me what or how to teach – I am the expert, not them.
30	<i>I will use the curriculum the teacher before me used at the schools where I will teach, even if it is ineffective, because I don’t want to cause any trouble.</i>
32	Generally, someone else decides what and how I teach.
33	In my physics class, I will combine textbooks and other materials, taking the best from each source.
35	The skills my students learn in my class, if any, will have little benefit to them once they graduate from school.
36	I find personal value in teaching.
37	The content I teach and the way that I teach it are not something for me to decide.
38	I feel responsible for doing my part in helping my students learn.
39	<i>I will not work more than my colleagues in preparing lessons, even if the quality of my lessons suffers.</i>
40	I think what my students learn in my class will be useful for them in other courses and/or in everyday life.
41	I think the progress of my students is independent of anything I as a teacher might do.
43	In some schools, teaching may suffer because the equipment in the physics classroom is outdated. There is nothing I can do about that as a teacher.
44	<i>I think the general skills my students learn in my class will give them a better chance at success in the future.</i>

*Note: Items that had been previously changed as a result of survey validation interviews are in italics.*

An expert panel (Wilson, 2004) consisting of 10 university faculty members at the UV who 1) teach and advise PSTs and 2) have experience teaching science in secondary school then assessed this finalized survey as a further measure of validation. These experts were provided the information and instruction in **Box 1** (in English):

**Box 1. Instructions provided to experts participating in the expert panel.**

This is a survey aiming to measure something we are calling “perceived agency” of a teacher or pre-service teacher, which we are defining as “a feeling of being in control over what is taught and of how it is taught”. This construct is similar to what Enghag and Niedderer (2008) describe as “ownership”, which “refers to the importance and need... to actually participate by discussions, choice, responsibility, and decision taking”. It also has a strong resemblance to what Milner-Bolotin (2001) described as “ownership of the learning process”. Milner-Bolotin’s subjects were learners in project-based-instruction group projects, and she described “ownership” in terms of three overlapping realms, with full overlap corresponding to most ownership: 1) does the learner find personal value in the project (e.g., finds the knowledge to be useful and/or connected to prior knowledge?); 2) does the learner feel in control to make decisions and be proactive?; and 3) does the learner feel responsible for the learning process and results of the project?

With this background in mind, please imagine a physics teacher who strongly believes that he/she is in control of what is taught and of how it is taught (that is, has a strong sense of “agency”). How would this physics teacher answer the following 44 questions? Please fill out this survey, not necessarily with your own answers or with the answers that you most want your pre-service teachers to select, but with what you think most reflects a sense of “agency”.

At least 8 out of the 10 experts agreed with our interpretations (for example, they chose “Agree” or “Strongly Agree” for Q.1) on 31 of the 44 questions. These 31 questions are in **Table 1**.

Finally, we administered the finalized survey to PSTs at UV (in German) and TGU (in Japanese) in the spring 2019 semester to further assess the validity and reliability of the instrument using Rasch analysis.

### *Rasch Analysis*

The finalized survey was administered at the start of the semester to PSTs at both UV (Feb. 2019, N=78, German) and at TGU (April 2019, N=66, Japanese) so as to collect pilot data, with which we could further assess reliability and validity of the instrument via Rasch analysis. To have a reasonable ratio of questions (31) to respondents, (e.g., p. 13 of SAS) we then combined these groups to treat them as coming from a single population (N=144). We will examine this assumption in our discussion below.

Whereas the common method of scoring an assessment is to sum the number of correct responses, Rasch analysis accounts for the fact that some items are more difficult than others, a judgment made by the model based upon how many respondents answer a given question

correctly. We fit the data collected from the finalized 31-item survey to a five-level Andrich rating scale model (RSM) (Andrich, 1978), a model in the Rasch family (Bond & Fox, 2007). While Rasch analysis is frequently used with dichotomous data (either the response is correct or it is not), the RSM is a frequently-used model for polytomous data, such as those obtained with Likert-scale surveys. Traditionally, surveys utilizing a Likert scale assume that the difference between a response of “3” and “4” on an item is equal to the difference between a response of “4” and “5” on the item (Likert, 1932). This is unfounded, however, as such labels on the responses are only ordinal data. Like with Rasch analysis in general, the RSM converts the ordinal data into ratio-scaled data, turning each response into a normalized score and then scoring each respondent accordingly. Furthermore, it also provides fit statistics on the items to indicate whether a given item is consistent with the instrument as a whole (namely, respondents who do better on the rest of the survey should do better on that item as well). It does this by fitting the data to a logistic equation in which the probability of choosing a particular response (“Neutral”, “Indicates agency”, etc.) is an exponential function of the difference between the respondents’ “trait” (level of perceived agency) and the “difficulty” of making that response. We performed Rasch analysis using WinSteps v.3.74 (Linacre, 2012b). Rasch models like the RSM have the following three specifications for data to be useful for the construction of the measurement (Wright, 1991):

- 1) There is only one dimension (here, “perceived agency”) that is assessed by the instrument, and one corresponding person measure (“how strong of a sense of agency does the respondent have?”).
- 2) As the person measure increases, so too does the likelihood of favorable responses on each item. Other than this trend, however, items are not correlated. That is, they are locally independent.
- 3) The model is a good fit to the data.

These three specifications should be investigated. First, we will discuss our findings regarding unidimensionality. It is important in calculating a total score for a respondent on the instrument. Namely, in summing scores across items to calculate a total score, it is implicitly assumed that all items measure the same construct. If, in fact, some items measure a different construct, then it is unclear what, if anything, a summative score indicates. At the same time, however, it has been argued that no instrument is perfectly unidimensional, and so the relevant question is rather whether or not the instrument would be improved by removal of items or by conceptualizing the instrument as a number of smaller instruments (p. 562, Linacre, 2012a). Rasch analysis assesses dimensionality via principal components analysis (PCA) of residuals. In looking at residuals, the analysis investigates the part of the data that is not accounted for by the model. If groups of items share the same patterns of unaccountability, then that may indicate a secondary dimension. In the analysis, the variance of the residuals (“unexplained variance”) is calculated, and the values are normalized such that each item contributes 1 eigenvalue to the total unexplained variance (p.

562, Linacre, 2012a). In the case of our data, the raw variance explained by the items was 19.1% and the persons explained another 11.5% of the raw variance, totaling 30.6% explained by the measures. Note that a small number explained by the measures like this is not automatic grounds for concern – indeed it is to be expected for cases when items are of comparable “difficulty” and persons are of similar “ability” (p. 561, Linacre, 2012a). However, more problematic was the finding that the first contrast (the largest secondary dimension) had an eigenvalue (an indicator of the total variation explained by that contrast) of 3.3. This is the strength of roughly three items. According to Linacre, “In the unexplained variance, a ‘secondary dimension’ must have the strength of at least 3 items, so if the first contrast has ‘units’ (i.e., eigenvalue) less than 3 (for a reasonable length test) then the test is probably unidimensional” (p. 561, Linacre, 2012a). The three items corresponding to our first contrast were identified by looking at clustering of items with extreme loadings on the contrast. These were Q.40, Q.18, and Q.35.

Q.40: *I think what my students learn in my class will be useful for them in other courses and/or in everyday life.*

Q.18: *What my students learn in my class will have little benefit for them in other courses and/or in everyday life.*

Q.35: *The skills my students learn in my class, if any, will have little benefit to them once they graduate from school.*

Together with Q.44, “I think the general skills my students learn in my class will give them a better chance at success in the future.”, we can indeed consider these items to form a secondary dimension on the survey. These four items were all based upon item 9 and item 12 of the Ownership Measurement Questionnaire (Milner-Bolotin, 2001): “I think I will be able to use what I learn in this project in other courses, or/and in everyday life” and “I think the skills I am learning in the project will help me to succeed in the future”, respectively. According to a Principal Component Analysis, Milner-Bolotin found that these items load heavily on a component corresponding to “personal value of the project.” In the case of our modification to these items, we can perhaps think of the secondary dimension we detected as “perceived value of instruction for student future.” We initially associated this dimension with our definition of “perceived agency” (“sense of being in control over what is taught and how it is taught”) with the justification that a teacher who does not sense the value of his or her instruction is unlikely to find motivation to exercise agency. However, we acknowledge that this connection is tenuous. For example, it is easy to imagine a teacher who recognizes the significance of the classroom on the future of a learner but, due, for example, to low sense of self-efficacy (which may particularly be the case for new teachers with relatively little experience, as discussed, for example, by (Yerdelen-Damar, Boz, & Aydın-Günbatır, 2017), decides to follow authority and not exercise agency. As such, we feel at present that these four items should be removed to bring the Perceived Agency Survey closer to being unidimensional and to ease interpretation of respondent score. Removing these four items decreased the eigenvalue of the first contrast below 3.0 (to 2.7) and

increased the raw variance explained by the measures marginally to 30.7%. Having made this improvement to the survey by removing these four items, we then assessed the other two model specifications with the resulting 27-item survey.

The specification of local independence was found to hold with the 27-item survey. The residual of a data point is the difference between the actual observation and what would be expected by the model. If the residuals of two items are correlated, that may indicate that the items are redundant or that they both measure some trait other than the main one targeted by the model. Hence, Rasch analysis looks for correlation between residuals of items as indication of lack of local independence. The largest standardized residual correlation was .45, between Q19 and Q39. This value is below .7, which is considered to be the cut-off for “highly locally dependent items” (p. 404, Linacre, 2012a).

Finally, we investigate the third specification, model fit. One should consider fit statistics at both the global level and at the level of the individual item. Looking first at item fit, we examined infit and outfit mean square values, which compare observed responses with those predicted by the model. Whereas infit is more sensitive to unexpected responses to items near the person’s measure level, outfit is more sensitive to responses on items much “easier” or “more difficult” than the level corresponding to the person’s measure (p. 411, Linacre, 2012a and p.165, Boone, Staver, & Yales, 2013). If these mean square values are too large, that indicates that the item is not as closely related to the overall construct as would be expected if the item measured the overall construct: that is, it likely measures something else. We found that all items exhibit mean square values of less than 2.0 logits (a ratio-level scale), which is considered to be a rule-of-thumb cut-off for acceptable fit. Namely, above this level, the off-variable noise becomes greater than the useful information provided by the item, resulting in the measurement being degraded (p. 166, Boone et al., 2013). Question 16 was the poorest-fitting item, with an outfit mean square value of 1.50.

In addition to fit statistics at the level of individual items, one should also inspect global fit statistics (Table 2). To this end, we calculated the Root Mean Square Error (RMSE). Data will not generally lie perfectly on the regression line created by the model (that is, there will be a prediction error, or “residual” for each data point). RMSE measures the standard deviation of these residuals. Although RMSE does not have a universal number as a cut-off below which the model is a “good” fit, it can be used between various best-candidate models, with the lowest value indicating the best-fitting model. The Perceived Agency Survey is a Likert-scale survey with the same five options for respondents to choose from in each item. The most obvious model to use is thus a five-category RSM. However, one can also imagine collapsing the survey into three categories, with a response of “Strongly Agree” being coded the same as a response of “Agree” (and likewise for “Strongly Disagree” with “Disagree”). This is commonly done with attitudinal surveys (e.g., (Adams et al., 2006; Redish et al., 1998). In addition, we considered also the possibility of using the Partial Credit Model (PCM) (Masters, 1982). The PCM is similar to the

RSM, but with an additional freedom. Suppose we have scored a response to an item with 3 points if the respondent selected “Neutral”, 4 points if the response (“Agree” or “Disagree”, depending on the prompt) indicates perceived agency, 5 points if the response (“Strongly Agree” or “Strongly Disagree”) strongly indicates perceived agency, etc. The RSM assumes that, across items, the difficulty for a respondent to choose 5 instead of 4 (for example) is a constant. The PCM does not have this restriction—the relative difficulty between levels can vary from item to item. As shown in Table 2, the lowest RMSE values were obtained when five levels were used, with the RSM having marginally smaller values than the PCM. Since the PCM is a more complex model than the RSM, there was little motivation to use any model other than the RSM with five levels.

**Table 2.** The global fit statistics.

Model	Person real RMSE (error)	Person model RMSE (error)	Item real RMSE (error)	Item model RMSE (error)
RSM 5-level	.29 (.59)	.26 (.60)	.11 (.49)	.11 (.50)
RSM 3-level	.48 (.54)	.47 (.55)	.12 (.29)	.11 (.29)
PCM 5-level	.30 (.61)	.26 (.62)	.11 (.37)	.11 (.37)
PCM 3-level	.50 (.58)	.48 (.60)	.10 (.24)	.10 (.24)

Note: Root Mean Square Error (RMSE) for various models. Rating scale model (RSM) 5-level demonstrated the smallest (and hence best) RMSE values, with Partial Credit Model (PCM) 5-level marginally higher. Since the RSM 5-level model is a more parsimonious model than the PCM 5-level model, the former was chosen as best-fitting.

### Validity and Reliability of the Instrument

Having attended to the three specifications made by the Rasch model, we can assess the validity and reliability of the instrument. Regarding validity, one measure is the person separation statistic. This is the ratio of the variance explained by the measures to the total variance (including error variance) (Bond & Fox, 2007). The 27-item Perceived Agency Survey exhibits a real person separation statistic of 2.06 and a model person separation statistic of 2.35, which are above the threshold of 2.0, indicating a “good level of separation” (p. 231, Boone et al., 2013). A further measure of validity is the observation that 4th-year PSTs at TGU (who had had experience teaching in schools prior to completing the survey) exhibited a relatively low level of perceived agency (0.66 logits (S.D. of .44), compared to an average of 1.27 logits (S.D. of .66) for the full data sample). This is consistent with literature on teacher training, which has indicated that teachers undergo “professional practice shock” when transitioning from the theory-heavy environment they experience as PSTs to their first years of in-service teaching (Mayring, 2014). Finally, although the 27 items are of comparable “difficulty”, there are some items that we would clearly expect to be more difficult to answer “correctly” than others. Namely, we made predictions for four items:

Q. 36: *I find personal value in teaching.* (We predicted that this item would be easy to agree with.)

Q. 14: *Teaching is just a job so I can get a paycheck – there is no benefit to me beyond that.* (Easy to disagree with)

Q. 23: *I will teach in the way I think is best, regardless of what my principal or other teachers might think.* (Hard to agree with)

Q. 2: *If the principal of my school tells me to teach in a certain way, I will do my best to teach that way, even if I don't really want to.* (Hard to disagree with)

Our predictions were largely met. Namely, the hardest three items were (in order), Q 4: *I prefer curriculum that tell the teacher exactly what to do, so that I don't risk making the wrong decision* (Apparently, this was extremely hard to disagree with), Q. 23, and Q. 2. Question 36 turned out to be a mid-range question, which is a surprising finding that warrants more attention (see Limitations and Future Development below). Question 14, on the other hand, was indeed the easiest item. The next two easiest items were Q. 21: *I think I have influence over the progress of my students* (easy to agree with) followed by Q. 38: *I feel responsible for doing my part in helping my students learn* (easy to agree with). In hindsight, it is not surprising that PSTs found these items to be quite easy.

In addition to the validity, we can also assess the reliability of the instrument. WinSteps calculated Cronbach's alpha to be .86. Cronbach's alpha is the most widely-used coefficient of internal consistency. This directly speaks to the reliability of the instrument, and indirectly to the unidimensionality of the instrument. Perhaps because of how widely it is used, this measure of reliability has received considerable criticism (e.g., Furr & Bacharach, 2008; Peters & Peters, 2014; Streiner & Norman, 2003). Cronbach's alpha is affected, for example, by local dependence between items as well as by the number of items in the questionnaire, such that merely adding more items to a survey can result in inflating alpha without any actual increase in internal coherency (Streiner & Norman, 2003). WinSteps also provides reliability measures based not upon the raw data (as Cronbach's alpha is), but rather upon the Rasch scores, which are linearized. These are the person reliability and item reliability measures, which measure the “reproducibility of relative measure locations”. When the reliability measures are high, it indicates that respondents and items estimated with relatively high (low) measures actually do have relatively high (low) measures (p. 644, Linacre, 2012a). The person reliability measures were .81 (real) and .85 (model). The item reliability measures were .95 (real) and .96 (model). All values are above .8, and are hence considered “good”. In the case of person reliability, being above the threshold of .8 indicates that the sample of respondents can be differentiated into 2-3 levels of perceived agency (low, medium, high) (p. 230, Boone et al., 2013).

We have preliminary data to speak to one final form of reliability, namely test-retest reliability. In measuring test-retest reliability, it is important to have the interval between tests be shorter than that expected for any actual change in the respondent to have occurred. It is also important, however, for the time interval to be long enough that the respondent does not remember his or her first set of responses (Wilson, 2004). As a result of being concurrently enrolled in two classes that completed the survey, three PSTs at UV were given the Perceived Agency Survey twice,

separated by more than one day but less than two weeks. These three respondents were quite consistent in their responses to the 27 survey items, changing by two levels (from “Agree” to “Disagree” or from “Neutral” to “Strongly Agree” or “Strongly Disagree”) an average of only 1.33 times, corresponding to 95% consistency.

## Discussion

This report has discussed our development and assessment of a survey to measure the perception of agency held by teachers, both in-service teachers and PSTs. Education researchers who study agency generally (e.g., Biesta et al., 2015; Lipponen & Kumpulainen, 2011; Olitsky, 2006; Smith, 2013) operate from within a sociocultural perspective (e.g., Lave & Wenger, 1991). Lipponen and Kumpulainen, for example, write “... human beings do not live in a vacuum... agency is not a fixed quality... not something people have. It is rather something that people do in social practice community. In sum, agency is interactive and cannot reside only in the individual because it is a socially constructed experience...” (Lipponen & Kumpulainen, 2011). Similarly, Biesta et al., write that “agency is understood as an emergent phenomenon of actor-situation transaction”, and, accordingly, utilize an ethnographic approach, interviewing only a few teachers and administrators about their teaching beliefs (Biesta et al., 2015). Although our study of perceived agency, which looks at isolated teachers outside of any relevant context (namely, taking a survey, as opposed to looking at their actions in the classroom), is in this regard unusual, we do not see that as deleterious. Surveys like ours are frequently-used tools by researchers operating from within a cognitivist framework. Greeno, writing from the situated/sociocultural perspective, contrasts the two frameworks: “The cognitive perspective takes the theory of individual cognition as its basis... The situative perspectives takes the theory of social and ecological interaction as its basis... While I believe that the situative framework is more promising, the best strategy for the field is for both perspectives to be developed energetically” (Greeno, 1997).

The benefit of a survey is that, in principle, data from a large number of respondents can be obtained with relative ease. We do, however, agree that greater depth of exploration is afforded by qualitative methods like interviewing, and we feel that a thorough study of agency cannot be limited to survey results. Firstly, surveys do not describe what the respondent actually does in practice, which is, ultimately, the topic of interest. For this reason, our survey results can only describe the respondents’ perceived agency, not the agency that they actually enact. Nor can we assume that a teacher will make the same statements on the survey as he or she would make talking to a friend during lunch, as surveys inherently do not account for the context-sensitivity of the respondents’ views. It is easy to imagine a teacher answering a question like “Do you feel like you have control over the learning of your students?” differently just before and just after a class, particularly if a carefully-planned lesson turns out to be a disaster. As with any survey, it is surely the case that our instrument introduces situations to some teachers that they had previously not considered, having a strong effect on the very thing we are aiming to measure. Whereas a score on a survey tends to suggest a fixed quality about the respondent (and consistency in

responses when taking the same survey twice is measured as a form of instrument reliability), ample evidence shows that people do not generally behave consistently. Barton et al. (Barton & Tan, 2010; Basu & Barton, 2009), in looking at students exercising agency in learning physics in an after-school program at the Boys and Girls Club, observed students who in some situations conformed “to the norm of passive learner or troubled student” but in other situations took on the role of, for example, “robotics expert”, depending on the task at hand and with whom the students were interacting (Basu & Barton, 2009). In the same way, we should expect teachers to exercise and perceive varying degrees of agency, depending on factors such as how cooperative the students are being, whether or not the teacher is being observed by the principal, and what lesson the teacher is teaching. As Davis and Krajcik write, “teachers’ use of and learning from text-based curriculum materials depend not only on the characteristics of the curriculum materials but also on the type of teaching activity in which the teacher is engaged” (Davis & Krajcik, 2005). Hence, whereas teachers may exercise agency in proactively learning from curriculum materials for some lessons, they might relax that agency and take a „business as usual“ approach for other lessons. Subtlety such as this cannot be captured by a survey. Despite these inherent limitations, surveys allow (limited!) insight into what is in the mind of the respondent, and that, in turn, plays a (limited!) role in what actually happens in the classroom.

It should be pointed out that, unlike many other attitudinal surveys (e.g., Adams et al., 2006; Redish et al., 1998), a “perfect score” on the Perceived Agency Survey is not necessarily something desired by educators of PSTs. One might argue that, regarding, for example, Q.29: “Parents should not tell me what or how to teach – I am the expert, not them”, it is undesirable to either “strongly agree” or “strongly disagree”, as a teacher should consider the wishes and suggestions of parents while still drawing upon his or her own expertise as a teacher. Davis and Krajcik describe this consideration by calling attention to the importance of balancing “prescription and autonomy” (Davis & Krajcik, 2005). We view the quantitative measurement of perceived agency then not as in setting a goal that “more is better”, but rather as an important step in describing an attribute relevant for teacher actions in the classroom.

## Limitations and Future Work

We assessed both the validity and reliability of the Perceived Agency Survey. Regarding validity, the 31 items in **Table 1** are the result of both survey validation interviews and an expert panel. However, we conducted both of these exclusively with the German version of the survey at one institution (UV). Future work could involve using the Japanese translation of the survey for comparable testing in Japan. Similarly, if the survey is translated into other languages, validity should be checked with interviewees and experts in those respective countries. We further assessed reliability and validity via Rasch analysis, which led to the removal of four additional items, Q.18, Q.35, Q.40, and Q.44, resulting in a 27-item survey.

The Person-Item Map (see **Figure 1**) shows that the sample of PSTs surveyed are not ideal for demonstrating the validity and reliability of the Perceived Agency Survey. Namely, the average



level of perceived agency in the respondents was significantly larger than that necessary to answer the average item “correctly”: the mean person measure was 1.27 logits greater than the mean item measure (two standard deviations). Several items had responses (particularly those pertaining to “Indicates a strong lack of agency”) that were chosen by NO persons.

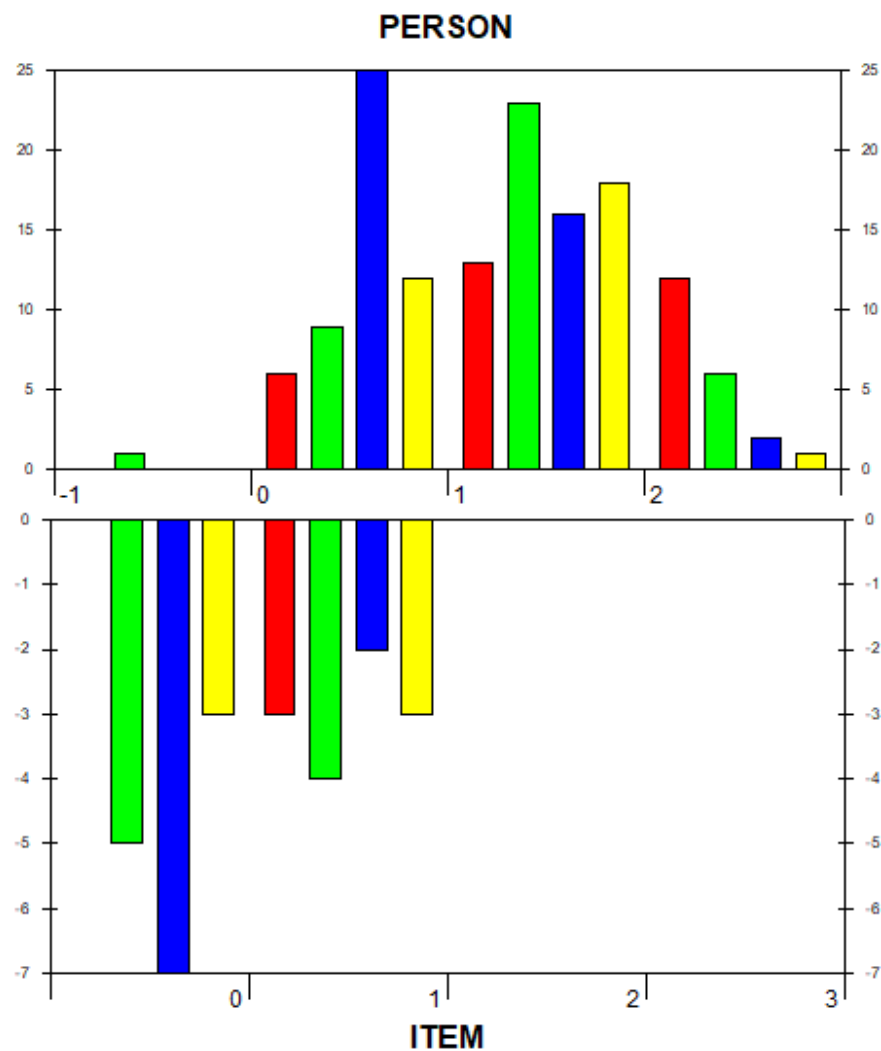


Figure 1. Person-Item Map of the 27-Item Perceived Agency Survey.

The average person measure was 1.27 logits, 2 standard deviations higher than the average item measure. Future development of the survey might involve removing additional items and replacing them with more “difficult” questions.

Furthermore, there continues to be concern regarding dimensionality. With the removal of the four items pertaining to “perceived value of instruction for student future,” the eigenvalue for the first contrast decreased from 3.3 to 2.7. As this is now below 3.0, Linacre comments, the test is “probably unidimensional” (p. 561, Linacre, 2012a). At the same time, however, a more common standard threshold for unidimensionality is for the eigenvalues of all contrasts to be less than 2.0. It might be fruitful to consider removing additional questions and replacing them with other items that are difficult, while still directly relating to the construct definition of “feeling in control over what is taught and how it is taught”. On the other hand, we think this survey could be of interest to researchers studying in-service physics teachers. While the responses of the PSTs on this survey generally indicated that the survey was too “easy” for them, in-service teachers might be a better match for this survey in general, as a result of “professional practice shock” (Mayring, 2014). The issue of unidimensionality should be revisited once data has been accumulated from in-service teachers. An additional opportunity for future exploration is the surprising result that Q. 36: I find personal value in teaching, turned out to be mid-level in difficulty. Additional survey validation interviews might shed light on why more PSTs than expected disagreed with this statement. If it is the case that the item is being misinterpreted, then that would make the item a candidate for removal or, at least, revision.

Finally, our decision to group the PSTs from UV and from TGU to consider them as coming from one population is questionable. Certainly, if the survey is to be used for international comparison, which we think would be an interesting next step, the validity and reliability should be assessed with respondents collected from diverse institutions within a given country. A similar consideration applies to our decision to collect responses from PSTs at various stages of their development at both UV and TGU. Namely, at UV, the 78 PSTs consisted of 23 PSTs from the seminar course taught by the first author, 36 PSTs at the start of first-semester conceptual lab, and 22 PSTs at the start of second-semester conceptual lab (three of these PSTs were in both the seminar and in one of the lab courses, and so only the first set of responses was used in the Rasch analysis). At TGU, the 66 PSTs consisted of 30 first-year students, 14 third-year students, and 22 fourth-year students, with varying degrees of teaching experience. Namely, only the fourth-year students had had practice teaching. Should the survey be used to assess the effects of instruction, which we also think would be an interesting future application, then care should be taken to accumulate a substantial pool of respondents pre-instruction and to treat them as a different population than respondents post-instruction. We wish to repeat here, that we do not interpret the respondent score on the Perceived Agency Survey as “more is better”. We can imagine excessively high scores correlating, for example, with arrogance and closed-mindedness, whereas the ideal teacher, one could argue, should be responsive to student and parent comfort and

contribute to a supportive and respectful school environment. Strongly agreeing with Q.29 or Q.23, “Parents should not tell me what or how to teach – I am the expert, not them” and “I will teach in the way I think is best, regardless of what my principal or other teachers might think” may indicate attitudes that compromise these goals. At the same time, however, increased agency does seem to be a target of some educators of both PSTs (e.g., Engle & Faux, 2006) and in-service teachers (e.g., Severance et al., 2016; Turpen, Olmstead, & Jardine, 2016). Rather than just giving teachers reformed curriculum and expecting them to implement it as faithful automatons, it is considered by many to be superior to have teachers reflect actively and agentively on the curriculum they are learning, perhaps even contributing to the creation of that curriculum (Severance et al., 2016). It might be interesting to see how respondents answer questions differently before and after such instruction.

## Conclusion

We created a survey to measure a “teacher’s perceived agency”, which we define as “a feeling of being in control over what is taught and of how it is taught”. Five survey validation interviews and a panel consisting of 10 experts suggest that 31 items on the survey are valid. Responses from 3 PSTs who took the survey twice indicate that the instrument is also reliable. Rasch analysis results from 144 PSTs further support the argument that 27 of those items are valid and reliable in distinguishing respondents based upon their level of perceived agency. Hence, we feel that accumulation of additional data with the survey is warranted. We welcome collaboration with additional instructors of PSTs and/or access to in-service teachers, and encourage those interested to contact the first author by e-mail. Although our survey is specific to physics, the questions themselves can readily be modified for use in other contexts, often by a change as trivial as replacing the word “physics” with “math”, etc. However, if researchers do make these or other changes to the survey, we recommend conducting survey validation interviews with the modified survey to ensure that the questions are understood as intended.

## Declaration of Conflict of Interests

The authors have no competing interests to declare.

## References

- Adams, W. K., Perkins, K. K., Podolefsky, N. S., Dubson, M., Finkelstein, N. D., & Wieman, C. E. (2006). New instrument for measuring student beliefs about physics and learning physics: The Colorado Learning Attitudes about Science Survey. *Physical Review Special Topics - Physics Education Research*, 2(010101), 1–14. <https://doi.org/10.1103/PhysRevSTPER.2.010101>
- Andrich, D. (1978). A rating formulation for ordered response categories. *Psychometrika*, 43, 561–573.
- Archer, L., DeWitt, J., Osborne, J., Dillon, J., Willis, B., & Wong, B. (2012). Science Aspirations, Capital, and Family Habitus: How Families Shape Children’s Engagement and Identification With Science. *American Educational Research Journal*, 49(5), 881–908.
- Ball, D. L., & Cohen, D. K. (1996). Reform by the Book: What Is—or Might Be—the Role of Curriculum Materials in Teacher Learning and Instructional Reform? *Educational Researcher*, 25(9), 6–14. <https://doi.org/10.3102/0013189X025009006>
- Bao, L., Cai, T., Koenig, K., Fang, K., Han, J., Wang, J., ... Wu, N. (2009, January 30). Physics: Learning and scientific reasoning. *Science*, Vol. 323, pp. 586–587. <https://doi.org/10.1126/science.1167740>
- Barton, A. C., & Tan, E. (2010). ‘We Be Burnin’! Agency, Identity, and Science Learning. *Journal of the Learning Sciences*, 19(2), 187–229.
- Basu, S. J., & Barton, A. C. (2009). Critical physics agency: further unraveling the intersections of subject matter knowledge, learning, and taking action. *Cult Stud of Sci Educ*, 4, 387–392.
- Biesta, G., Priestley, M., & Robinson, S. (2015). The role of beliefs in teacher agency. *Teachers and Teaching*, 21(6), 624–640. <https://doi.org/10.1080/13540602.2015.1044325>
- Bond, T., & Fox, C. (2007). *Applying the Rasch Model. Fundamental Measurements in the Human Sciences* (2nd ed.). Mahwah, NJ: Erlbaum.
- Boone, W. J., Staver, J. R., & Yales, M. S. (2013). *Rasch analysis in the human sciences*. Springer Science & Business Media.
- Davis, E. A., & Krajcik, J. S. (2005). Designing Educative Curriculum Materials to Promote Teacher Learning. *Educational Researcher*, 34(3), 3–14. <https://doi.org/10.3102/0013189X034003003>
- Davydov, V. V., Slobodchikov, V. I., & Tsukerman, G. A. (2003). The Elementary School Student as an Agent of Learning Activity. *Journal of Russian & East European Psychology*, 41(5), 63–76.
- DeBarger, A. H., Penuel, W. R., Moorthy, S., Beauvineau, Y., Kennedy, C. A., & Boscardin, C. K. (2017). Investigating Purposeful Science Curriculum Adaptation as a Strategy to Improve Teaching and Learning. *Science Education*, 101(1), 66–98. <https://doi.org/10.1002/sc.21249>
- Edwards, A., & D’arcy, C. (2004). Relational agency and disposition in sociocultural accounts of learning to teach. *Educational Review*, 56(2), 147–155.
- Enghag, M., & Niedderer, H. (2008). Two Dimensions of Student Ownership of Learning During Small-Group Work in Physics. *International Journal of Science and Mathematics Education*, 6(4), 629–653. <https://doi.org/10.1007/s10763-007-9075-x>
- Engle, R. A., & Faux, R. B. (2006). Towards Productive Disciplinary Engagement of Prospective Teachers in Educational Psychology: Comparing Two Methods of Case-based Instruction. *Teaching Educational Psychology*, 1(2).
- Furr, R. M., & Bacharach, V. R. (2008). *Psychometrics*. Thousand Oaks, CA: Sage.
- Greeno, J. G. (1997). On Claims That Answer the Wrong Questions. *Educational Researcher*. <https://doi.org/10.3102/0013189x026001005>
- Hake, R. R. (1998). Interactive-engagement versus traditional methods: A six-thousand-student survey of mechanics test data for introductory physics courses. *American Journal of Physics*, 66, 64–74.
- Harlow, D. B. (2009). Structures and improvisation for inquiry-based science instruction: A teacher’s adaptation of a model of magnetism activity. *Science Education*, 142. <https://doi.org/10.1002/sc.20348>
- Harlow, D. B. (2010). Uncovering the Hidden Decisions that Shape Curricula. *AIP Conference Proceedings*, 21–24.
- Hull, M. M. (2013). *Do students have cultural scripts? Results from the first implementation of open source tutorials in Japan*. Unpublished doctoral dissertation. University of Maryland, College Park.
- Hull, M. M., & Uematsu, H. (2020). Development of Preservice Teachers’ Sense of Agency. *Research and Innovation in Physics Education: Two Sides of the Same Coin*. Springer, Cham, 209–223.
- Hull, M. M., Vormayr, K., & Uematsu, H. (2019). Validation of a Survey to Measure Pre-service Teachers’ Sense of Agency. *Proceedings of GIREP Conference*.
- Lave, J., & Wenger, E. (1991). *Situated learning: Legitimate peripheral participation*. Retrieved from

- <https://books.google.at/books?hl=de&lr=&id=CAVIOrW3vYAC&oi=fnd&pg=PA11&ots=OCqAyqXFfi&sig=ej03O-oTDJPFvjBtSs0LqwdtFNM>
- Likert, R. A. (1932). A technique for the measurement of attitudes. *Archives of Psychology*, 140, 1–55.
- Linacre, J. M. (2012a). *A User's Guide to WINSTEP® MINI STEP Rasch-Model Computer Programs Program Manual 3.74.0*.
- Linacre, J. M. (2012b). *Winsteps® Rasch measurement computer program*. Retrieved from Winsteps.com
- Lipponen, L., & Kumpulainen, K. (2011). Acting as accountable authors: Creating interactional spaces for agency work in teacher education. *Teaching and Teacher Education*, 27(5), 812–819. Retrieved from <https://www.sciencedirect.com/science/article/pii/S0742051X11000023>
- Masters, G. N. (1982). A Rasch model for partial credit scoring. *Psychometrika*, 47, 149–174.
- Mayring, P. (2014). *Qualitative content analysis: theoretical foundation, basic procedures and software solution*.
- Milner-Bolotin, M. (2001). *The effects of topic choice in project-based instruction on undergraduate physical science students' interest, ownership, and motivation* (University of Texas at Austin). Retrieved from [https://www.researchgate.net/profile/Marina\\_Milner-Bolotin/publication/35219406\\_The\\_effects\\_of\\_topic\\_choice\\_in\\_project-based\\_instruction\\_on\\_undergraduate\\_physical\\_science\\_students'\\_interest\\_ownership\\_and\\_motivation/links/5769c34608ae2d7145ba86c0.pdf](https://www.researchgate.net/profile/Marina_Milner-Bolotin/publication/35219406_The_effects_of_topic_choice_in_project-based_instruction_on_undergraduate_physical_science_students'_interest_ownership_and_motivation/links/5769c34608ae2d7145ba86c0.pdf)
- Olitsky, S. (2006). Structure, agency, and the development of students' identities as learners. *Cultural Studies of Science Education*, 1(4), 745–766. <https://doi.org/10.1007/s11422-006-9033-x>
- Perceived Choice and Awareness of Self Scale (PCASS). (2019). Retrieved September 28, 2018, from <http://selfdeterminationtheory.org/pcass/>
- Peters, G.-J., & Peters, G.-J. Y. (2014). The alpha and the omega of scale reliability and validity Why and how to abandon Cronbach's alpha and the route towards more comprehensive assessment of scale quality. *European Health Psychologist*. <https://doi.org/10.31234/osf.io/h47fv>
- Redish, E. F., Saul, J. M., & Steinberg, R. N. (1998). Student expectations in introductory physics. *American Journal of Physics*. <https://doi.org/10.1119/1.18847>
- Robertson, A. D. (2018). Supporting the development of curricular knowledge among novice physics instructors. *American Journal of Physics*, 86(4), 305–315.
- SAS, S. (n.d.). *PRINCIPAL COMPONENT ANALYSIS*. Retrieved from <https://support.sas.com/publishing/pubcat/chaps/55129.pdf>
- Sawyer, R. K. (2004). Creative Teaching: Collaborative Discussion as Disciplined Improvisation. *Educational Researcher*, 33(2), 12–20. <https://doi.org/10.3102/0013189X033002012>
- Schrittmesser, I. (2013). *From Novice to Professional: Teachers for the 21st century and how they learn their job. Learning to Be a Teacher in a Changing World*. University of Barcelona.
- Severance, S., Penuel, W. R., Sumner, T., & Leary, H. (2016). Organizing for Teacher Agency in Curricular Co-Design. *Journal of the Learning Sciences*, 25(4), 531–564. <https://doi.org/10.1080/10508406.2016.1207541>
- Sewell, W. H. (1992). A Theory of Structure: Duality, Agency, and Transformation. *American Journal of Sociology*, 98(1), 1–29. <https://doi.org/10.1086/229967>
- Smith, J. A. (2013). Structure, Agency, Complexity Theory and Interdisciplinary Research in Education Studies. *Policy Futures in Education*, 11(5), 564–574. <https://doi.org/10.2304/pfie.2013.11.5.564>
- Streiner, D. L., & Norman, G. R. (2003). *Health measurement scales: A practical guide to their development and use* (3rd ed.). New York: Oxford University Press Inc.
- Tosa, S. (2011). Comparing US and Japanese Inquiry-Based Science Practices in Middle Schools. *Middle Grades Research Journal*, 6(1), 29–46.
- Wei, B., & Chen, N. (2019). Agency at work: two beginning science teachers' stories in a context of curriculum reform in China. *International Journal of Science Education*. <https://doi.org/10.1080/09500693.2019.1600205>
- Wilson, M. (2004). *Constructing measures: An item response modeling approach*. Retrieved from <https://content.taylorfrancis.com/books/download?dac=C2010-0-35526-4&isbn=9781135618056&format=googlePreviewPdf>
- Wright, B. (1991). Scores, Reliabilities and Assumptions. *Rasch Measurement Transactions*, 5(3), 157–158. Retrieved from <https://www.rasch.org/rmt/rmt53a.htm>
- Yerdelen-Damar, S., Boz, Y., & Aydın-Günbatar, S. (2017). Mediated Effects of Technology Competencies and Experiences on Relations among Attitudes Towards Technology Use, Technology Ownership, and Self Efficacy about Technological Pedagogical Content Knowledge. *J. Sci Educ Technol*, 26, 394–405.

