Validation of a survey to measure pre-service teachers' sense of agency

Michael M HULL, Katharina VORMAYR and Haruko UEMATSU

1University of Vienna, Austrian Education Competence Centre Physics, Porzellanasse 4, 1090, Vienna, Austria
2Tokyo Gakugei University, Department of Physics, Nukuikita-machi 4-1-1, Koganeishi, 184-0015, Tokyo, Japan

Abstract. Dissemination of reformed curriculum requires teachers to feel that they have the freedom to implement the curriculum in the classroom. Even instructors who are trained in research-based instruction and are convinced of its value might fail to implement the curriculum in the classroom if, for example, they feel a lack of support from their school or colleagues. In this talk, we will report on the validation process of a survey designed to measure a “teacher's or pre-service teacher's perceived agency”, which we define as “a feeling of being in control over what is taught and of how it is taught.”

1. Introduction

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 instructors themselves. Many physics instructors 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 instructors who are convinced of the value of research-based instruction and are trained in it might fail to implement the curriculum in the classroom if their school or colleagues do not support their attempts. In Japan, for example, it is not uncommon for a school to resist change, as a result of inertia. If a teacher feels that colleagues or even the principal is hostile towards the reformed instruction in which the teacher was trained in, then this will naturally hinder the dissemination of the new material. We argue, however, that it is not the control that the teacher does or does not have that directly 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 instructor did, then the reformed curriculum will still not be disseminated. We define “a teacher’s or pre-service teacher’s perceived agency” as “a feeling of being in control over what is taught and of how it is taught” [1]. In this report, we will discuss the validation process of a survey designed to gain insight into this trait.

2. Motivation

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. [2]), their attitudes about the nature of physics knowledge and learning tend to get worse (e.g., [3]), 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 content knowledge, such as Maxwell's equations and Newton's laws of motion, PSTs are generally taught 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. 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 [4].

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. Other teachers may hesitate to use reformed curriculum out of consideration of the status quo. Anecdotally, 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. We find the construct of “agency” to be useful in describing these kinds of hesitations that teachers might exhibit. We do not mean “agency” in the sense of being an extension of an organization, like a secret agent. Rather, our usage is consistent with the definition of “the capacity to initiate purposeful action that implies will, autonomy, freedom, and choice” [5]. We 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 [4,6,7], colleagues at the school [4], or by other factors including student [4,6] and parent [6] expectations.

In cases like those described above, 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. Rather, there is 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 [8–10]. Accordingly, it should be expected that two teachers in the same school would 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, we see the question of how these structures influence the views of teachers to also be significant, as it is these views that will more directly lead to action in the classroom. Hence, rather than studying the “capacity” of a teacher or PST “to initiate purposeful action that implies will, autonomy, freedom, and choice” (p. 813 of [5]), we focus on how PSTs perceive this capacity. The research questions underlying our work are “is perceived teacher/PST agency a construct that can be measured by a survey?”, and “is such a survey useful?”

3. Perceived Agency Survey Design

As discussed in detail in [1], the Perceived Agency Survey was inspired by two existing surveys which measure agency in other contexts. 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 [11]. 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 [12]. As an example, one of the prompts from the Ownership Measurement Questionnaire reads as follows: “...
Questionnaire was “I feel responsible for the in-depth exploration of the project-topic” (p. 172 of [11]), which is similar to Question 42 (Q.42) of our Perceived Agency Survey: “I feel responsible for making my students think deeply.” Other prompts were added based upon common teaching experiences, such as Q.8: “I need to listen carefully to the demands of the parents of my students to make sure I’m teaching what they want their children to learn.” The original survey design utilized a paired-prompt format. For example, Q.8 is paired with Q.29: “Parents should not tell me what or how to teach – I am the expert, not them.” In total, we constructed 44 5-level Likert-scale items. As is commonly done in analysis of attitudinal surveys (e.g., [3,13]) we collapsed “Strongly Agree” and “Agree” (and likewise “Strongly Disagree” and “Disagree”) into one code.† Approximately half of the items were reverse-coded, as “Disagree” indicated a perception of agency for those items. Respondents were told to leave an item blank if they do not understand the statement and to select “Neutral” if they have no opinion about the statement.

It should be pointed out that, unlike many other attitudinal surveys (e.g., [3,13]), a “perfect score” on the agency survey is not necessarily something desired by educators of PSTs. One might argue that, regarding Q.8 and Q.29, 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. 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.

4. Survey Validity and Reliability

The first and third authors created the Perceived Agency Survey first in English and then in Japanese for pilot administration to PSTs both at UV and at TGU in the spring 2018 semester. The second author, a native German speaker, translated the survey into German in the fall 2018 semester, and then used it to conduct a succession of survey validation interviews which led to a finalized survey. An expert panel [14] then assessed this finalized survey as a further measure of validation. Furthermore, we administered the finalized survey to PSTs at UV (in German) and TGU (in Japanese) in the spring 2019 semester to further assess the construct validity of the questionnaire via factor analysis, as well as to check the reliability of the instrument using Cronbach's alpha.

4.1. Survey validation interviews

At the start of the fall 2018 semester, the second author visited the classroom of the first author at UV and the first author left the classroom while the second author invited PSTs to participate in survey validation interviews. The second author assured the PSTs of anonymity, and five PSTs agreed to partake in the interview. The second author conducted interviews with each of these five PSTs in German. These interviews were in the think-aloud style (aka “cognitive labs”), where the interviewee talks aloud about what he/she is thinking in real time while responding to each item on the survey in turn. In cases of ambiguity, the interviewer asks follow-up questions. In addition, if it seems that the interviewee is misinterpreting the intention of a question, the interviewer rephrases the question, or even asks the interviewee to suggest how to reword the question. As an example, the following exchange (translated from German) took place during the third interview when the respondent reached Q.10, which at that point in development read “From time to time, because of the nationally-required curriculum, physics teachers have to teach in a way that is not very effective.”

PST: There I would pretty much choose a 3 [Neutral]. I would say that certain things are maybe not 100% optimal. But in many cases, the curriculum is such that you can set it up exactly how you yourself think will result in good teaching, to optimize it.

† Note that this practice is not universally accepted, as it “flattens” the differences between responses. Although we have not investigated the effects of maintaining all five categories in this report, we intend to do so in future work.
Interviewer: *That means that you are thinking about the things that the curriculum makes you teach?*

PST: *Exactly.*

Interviewer: *OK, because I would understand that the question is talking about the curriculum telling you how you should teach and not what. Would it change your answer, if the question were only talking about how?*

PST: *Well, then I would maybe change to a 2 [Disagree], because the requirements are very loose, and therefore don’t restrict you.*

Here, the interviewer interpreted the PST’s response regarding “certain things” to potentially refer to topics that are mandated by the curriculum. Since the intent of the question was not to ask about these topics, but rather about the means of teaching those topics, the interviewer asked the follow-up question regarding whether or not the PST’s response would change if considering not the topics, but the teaching approach. After each interview, the interviewer met with the first author to discuss suggestions for modification to the survey in light of the interview (without revealing the name of the PST interviewed, of course), considering transcript such as that just presented. Changes were often subsequently made to the survey, and the revised survey was then used for the next survey validation interview. The researchers repeated this cycle five times throughout the fall 2018 semester, culminating in a finalized survey. This 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.

4.2. Expert panel

After the five survey validation interviews, we presented the finalized survey to an expert panel consisting of 10 university faculty members at the UV who 1) teach and advise PSTs and 2) have experience teaching science in secondary school. After providing some background theoretical information about how we are defining “perceived agency” as well as the history of the construct, we asked the experts to “…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.

4.3. Factor 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. This data collection served two purposes. First, it allowed us to carry out a factor analysis, which would serve as a further test of construct validity. Second, we used PST responses on the survey to calculate Cronbach’s alpha, which served as a measure of reliability. To have a reasonable ratio of questions (31) to respondents, (e.g., p. 13 of [15]) 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.

To explore to what extent the items on the survey hung together under one over-arching construct (perceived agency), a factor analysis was performed. First, a principal component analysis (PCA) was carried out to identify the number of factors. Since each variable is on the same Likert scale, the covariance matrix was used with the princomp function in R. In a PCA, a missing response invalidates the data (p. 13 of [15]), and so these respondents were removed from the data set. This resulted in 142 respondents. There are several criteria to consider when determining how many factors are significant. First, consistent with the Kaiser criterion [16], the standard deviations produced by R for each factor were squared to calculate the variance accounted for by the factor. As each item on the survey contributes a variance of 1 to the total variance, only factors accounting for a variance greater than 1 are taken to be meaningful. The variances accounted for by the first three factors were 1.66, 0.90, and 0.76. Using this criteria, only the first factor is significant. This was confirmed by a scree plot (Figure 1),
where a large drop in variance occurs from the first to second factor and subsequent drops are all of about the same size.

One should also look at the proportion of variance accounted for by each factor (good to be at least 5% to 10%) as well as the cumulative proportion of variance explained (good to be at least 70%). These values, as well as the standard deviations for each of the first 11 factors, are presented in the Table 2 below. As can be seen, each of the first six factors explain 5% or more of the variance.

### Table 1. The 31 questions remaining after the expert panel. Items that had been previously changed as a result of survey validation interviews are in italics. A * after the item number indicates that a response of “(Strongly) Disagree” was coded as reflecting perceived agency.

|   | Description                                                                                                                                 |
|---|--------------------------------------------------------------------------------------------------------------------------------------------|
| 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 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.   |
| 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 | 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. |
| 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 think I have influence over the progress of my students.                                                                                   |
| 23| I will teach in the way I think is best, regardless of what my principal or other teachers might think.                                     |
| 27| Curricular resources are, at most, a guide for teachers to use or modify creatively, as the situation requires.                           |
| 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 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. |
| 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 will not work more than my colleagues in preparing lessons, even if the quality of my lessons suffers.
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 think the general skills my students learn in my class will give them a better chance at success in the future.

![Scree plot indicating additional variance accounted for by each additional factor ("component"). The drop from 1 to 2 factors is notably larger than between other factors.](image)

**Figure 1.**

**Table 2.** Results for the first 11 factors from principal component analysis.

| Factor | 1    | 2    | 3    | 4    | 5    | 6    | 7    | 8    | 9    | 10   | 11   |
|--------|------|------|------|------|------|------|------|------|------|------|------|
| S.D.   | 1,293| 0,949| 0,876| 0,817| 0,786| 0,756| 0,699| 0,665| 0,630| 0,598| 0,588|
| P.V.   | 0,165| 0,089| 0,076| 0,066| 0,061| 0,056| 0,048| 0,044| 0,039| 0,035| 0,034|
| C.P.   | 0,165| 0,254| 0,330| 0,396| 0,457| 0,513| 0,562| 0,605| 0,644| 0,680| 0,714|

* Standard deviation.
* Proportion of variance.
* Cumulative proportion.

From here, a factor analysis was performed to describe the six factors that explain at least 5% of the variance each. This analysis utilized a varimax rotation, such that factors would be uncorrelated to each other. The loadings of the 31 items on these six factors is shown in Table 3. Loadings that are “large” (greater in absolute value to 0.4, (p. 29 of [15]) are underlined. We see that Q.18, Q.35, Q.40, and Q.44 load strongly on factor 1. These are 4 items that one would expect to all measure the same thing: how useful is the lesson for students outside of the classroom? These results, together with the other five factors and the items that load strongly on them are in the central column of Table 3. Factor 3 consists of 2 items that discuss working more than colleagues if needed. Factor 5 consists of 2 items that discuss being in control over what is taught and how it is taught. Factors 2 and 4, however, show some unexpected results. Namely, Q.30 in factor 2: “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”, we would intuit, should produce responses that correlate with those from Q.2 in factor 4: “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.” Looking at the raw data, however, we see that this was not particularly the case. Question 30, in fact, was quite easy for PSTs to disagree with. Of the 144 PSTs, 91% either disagreed or strongly
disagreed with the statement. In contrast, only 52% of PSTs disagreed with Q.2. Similarly, we might expect that Q.33 in factor 2, “In my physics class, I will combine textbooks and other materials, taking the best from each source” is equivalent to Q.20 in factor 4, “Once I choose a physics textbook for my classroom, I will follow it carefully.” However, 90% of PSTs agreed with Q.33, whereas only 63% of PSTs disagreed with Q.20. In hindsight of the data, these discrepancies are perhaps not so surprising. It is quite a bit different to cause trouble to colleagues than to ignore the requests of a boss. We might thus consider that factor 4 indicates a PSTs inclination to follow instructions, be it from the principal or the curriculum/textbook. The item with the greatest loading in factor 2 is Q.38, “I feel responsible for doing my part in helping my students learn.”, and we can imagine a PST who strongly feels this way to modify curriculum accordingly (Q.3 and Q.33) and not use ineffective curriculum (Q.30), because the PST thinks that he or she has influence over the progress of students (Q.21).

With these six factors, only a total of 51% of the variance is accounted for. As we see in Table 2, keeping 11 factors allows for 70% of the variance to be accounted for, although this results in factors that individually account for only 3% of the variance. Furthermore, in addition to Q.30 and Q.2 being in different factors (like wise for Q.20 and Q.33), we see in the right column of Table 3 that Q.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” is not in the same factor as Q.30, as we would intuit. It is hard to justify conceptually why this should be the case. Looking at the raw data, 88% of respondents agreed with Q.9 (recall from just above that 91% of PSTs disagreed with Q.30), and 82% of respondents gave a consistent code across the two items (indicating agency, neutral, or indicating a lack of agency on BOTH items). Additional data is desirable to ascertain whether the survey is best analyzed in terms of 11 factors, 6 factors, or just 1 factor.

In particular, as shown in Figure 2 below, survey respondents tended to answer items with the response indicating perceived agency more often than not. We suspect that this led to some ceiling effects. Future work should involve administering the survey to in-service teachers as well, to see if their responses are as positive as those of these PSTs.

### Table 3. Loadings of the 31 items on the first six factors. The bottom-right row is Cronbach’s Alpha, calculated for each factor. Each item is assigned to the factor for which it has the greatest loading (Q.1 is in factor 2, for example).

|   | 1   | 2   | 3   | 4   | 5   | 6   |   | 1   | 2   | 3   | 4   | 5   | 6   |
|---|-----|-----|-----|-----|-----|-----|---|-----|-----|-----|-----|-----|-----|
| Q1| 0.37|     | 0.11| 0.12|     |     | Q27| 0.10| 0.11| 0.17| 0.13|     |     |
| Q2|     | 0.47| 0.14|     |     |     | Q28| 0.16| 0.35| 0.32|     |     |     |
| Q3| 0.11| 0.46| 0.24|     |     |     | Q29|     | 0.10| 0.15| 0.83|     |     |
| Q4| 0.13| 0.66|     |     |     |     | Q30| 0.19| 0.43| 0.31| 0.23|     |     |
| Q6| 0.29| 0.19| 0.27| 0.16|     |     | Q32| 0.13| 0.11| 0.79|     |     |     |
| Q9| 0.36| 0.10| 0.37|     |     |     | Q33| 0.32| 0.44| 0.22| 0.25|     |     |
| Q13| 0.13| 0.39| 0.14|     |     |     | Q35| 0.67| 0.12| 0.22| 0.15| 0.20|     |
| Q14| 0.30| 0.15| 0.30| 0.14| 0.13|     | Q36| 0.15| 0.24| 0.37|     | 0.11|     |
| Q15| 0.16| 0.13| 0.24| 0.25|     |     | Q37| 0.11| 0.42| 0.12| 0.14|     |     |
| Q16| 0.13| 0.31| 0.18| 0.13|     |     | Q38|     | 0.76| 0.18| 0.15| 0.26|     |
| Q17| 0.13|     | 0.17| 0.56|     |     | Q39| 0.20| 0.15| 0.83| 0.13| 0.24|     |
| Q18| 0.60| 0.32|     |     |     |     | Q40|     | 0.96| 0.23|     | 0.12|     |
| Q19|     |     | 0.80| 0.17|     |     | Q41|     | 0.36| 0.26|     |     |     |
| Q20| 0.12| 0.16|     | 0.50| 0.31| 0.21| Q43|     |     | 0.26| 0.10|     |     |
| Q21| 0.12| 0.54|     |     |     |     | Q44|     |     | 0.71| 0.23|     |     |
| Q23|     | 0.23| 0.24|     |     |     | CA|     | 0.82| 0.73| 0.63| 0.56| 0.56| 0.31|
4.4. Cronbach’s alpha as a measure of reliability

The measures discussed so far have addressed the construct validity of the Perceived Agency Survey. Namely, does the survey measure what we intend it to measure? It is also important to consider the reliability of an instrument, which is a measure of how likely the test is to produce the same results for two different measurements. One means of measuring this is test-retest reliability, where the same respondents are given the same instrument within a time interval 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.

Table 4. Items that load heavily on factors. The Kaiser criterion and scree plot suggest keeping just 1 factor (left column). To account for 70% of the variance, 11 factors are needed (right column). However, only six of those factors individually account for at least 5% of the variance (middle column).

| Factor 1 | Keep 1 Factor | Keep 6 Factors | Keep 11 Factors |
|----------|---------------|----------------|-----------------|
|          | 14, 18, 19, 30, 33, 35, 36, 38, 39, 40, 44 | 18, 35, 40, 44 | 35, 40, 44 |
| Factor 2 | - | 3, 21, 30, 33, 38 | 13, 21, 30, 33 |
| Factor 3 | - | 19, 39 | 19, 39 |
| Factor 4 | - | 2, 4, 20 | 17, 32 |
| Factor 5 | - | 17, 32 | 20 |
| Factor 6 | - | 29 | 28 |
| Factor 7 | - | - | 2, 23 |
| Factor 8 | - | - | 36 |
| Factor 9 | - | - | 15, 27 |
| Factor 10 | - | - | 9, 16 |
| Factor 11 | - | - | 29 |

Figure 2. Average response for each question for each group of students. Blue bars (left) indicate the mean for UV (N=78) and red bars (right) indicate the mean for TGU (N=66). Error bars indicate one standard deviation. These values were calculated by attributing 2 points to a response indicating perceived agency, 0 points to a response indicating a lack of perceived agency, and 1 point to a response of “Neutral”. Median values below the x-axis are taken from all five levels of the Likert scale, with a maximum value of 4 indicating the most perceived agency (and 0 indicating least).

That is, if we imagine brainwashing the respondent to forget the first set of responses, will the respondent give the same responses the second time around? An extreme usage of this “brainwashing analogy” is
Cronbach’s alpha, which measures how consistent the questions are with each other. In doing so, each item is treated as its own stand-alone instrument, and it is assumed that the respondent is not answering questions based upon responses to previous items. We used the alpha function in R with the full set of data (N=144, 31 items) to calculate an alpha of 0.81, which is considered “good”. When separating the instrument into six factors, on the other hand, alpha for individual factors were generally below 0.70, which is considered the threshold for “acceptable” (see bottom-right row of Table 3). Similarly, split-half alphas were also low, with 0.70 for the 15 items with a * in Table 1, and 0.72 for the remaining 16 items. It therefore seems most promising to treat the instrument as a single test, instead of two half-tests or six factors.

5. Discussion

5.1. Does it make sense to measure perceived agency with a survey?

Education researchers who study agency generally (e.g., [4,5,9,10]) operate from within a sociocultural perspective (e.g., [17]). 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…” [5].

Although our study of perceived agency, which looks at isolated teachers/PSTs outside of any relevant context (that is, 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.” [18]

Certainly, we do not pretend to be able to make any strong claims with our survey on what will actually happen in the classroom, which is, ultimately, the topic of interest. Nor can we assume that a PST or in-service teacher will make the same statements on the survey as he or she would make talking to a friend during lunch. As with any survey, it is surely the case that our instrument introduces situations to some PSTs that they had previously not considered, having a strong effect on the very thing we are aiming to measure. Nevertheless, 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.

5.2. Limitations and Future Work

We assessed both the construct validity and reliability of the Perceived Agency Survey. Regarding construct 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 validity via factor analysis, but the results at present are inconclusive. As discussed above, the respondents to whom we administered this survey are not an ideal match for the survey, as their responses to the survey items were predominantly positive, likely leading to ceiling effects. As such, we intend to administer the survey to additional groups of respondents, such as in-service teachers, who might demonstrate a weaker sense of agency than the PSTs did.

Regarding reliability, our assessment at present consists exclusively of calculating Cronbach’s alpha. Cronbach’s alpha has been criticized for a number of reasons (e.g., [19]), however, so other means of reliability assessment, such as calculating other comparable statistics, or test-retest reliability, should be investigated. We are particularly interested in fitting the data to a rating scale model, a model for polytomous data within the Rasch family, with which we could assess reliability. This would also enable
us to rank the items and to compare the ranking with our own expectations, which would serve as an additional measure of validity.

Finally, our decision to group the PSTs from UV and from TGU to consider them as coming from one population is questionable. Examination of Figure 2 shows that PSTs at TGU had lower averages than counterparts at UV with nearly every item. On Q.29, for example, the average UV score was 1.77, but only 1.08 at TGU. However, in addition to the significant overlap of error bars indicated in Figure 2, a cluster analysis of the respondents did not show noteworthy clustering of TGU PSTs vs UV PSTs. Hence, we feel that this decision to combine the data as we did is justified as a first step. With the accumulation of additional data, however, differences in groups may turn out to be statistically significant. Certainly, if the survey is to be used for international comparison, which we think would be an interesting next step, construct 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). 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. However, we felt justified in pooling these responses as the data spread was very large in comparison to differences in means. However, 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.

6. Implications and Conclusion
The research questions underlying our work are “is perceived teacher/PST agency (that is, “the feeling of being in control over what is taught and how it is taught”) a construct that can be measured by a survey?”, and “is such a survey useful?” To investigate the first research question, we created the Perceived Agency Survey and accumulated pilot data to investigate 1) its content validity via survey validation interviews, expert panel, and factor analysis, and 2) its reliability via calculation of Cronbach’s alpha. Although our results are only preliminary, we remain optimistic that the answer is “yes, perceived teacher/PST agency is a construct that can be measured by a survey.” Although we have not yet begun investigation of the second research question of whether such a survey is useful or not, we imagine that administration of a finalized version of the Perceived Agency Survey could have implications for PST/teacher educators; namely, the survey could be used to indicate effects of workshops for teachers or courses for PSTs, by administering the survey at the start and end of the course and looking at changes. Use of the survey to look for similarities and differences in teachers/PSTs in different countries could also potentially have implications for school administrators and developers of national education standards. In summary, we feel that accumulation of more data is warranted to further explore these research questions. 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.

7. References
[1] Hake RR 1998 American Journal of Physics 66 64-74
[2] Redish EF, Saul JM, and Steinberg RN 1998 American Journal of Physics 66 212-224
[3] Biesta G, Priestley M and Robinson S 2015 Teachers and Teaching 21 624-40
[4] Lipponen L and Kumpulainen K 2011 Teaching and Teacher Education 27 812-819
[5] Ball DL and Cohen DK 1996 Educational Researcher 25 6-14
[6] Wei B and Chen N 2019 International Journal of Science Education 1-16
[7] Sewell WH 1992 American Journal of Sociology 98 1-29
[8] Smith JA 2013 Policy Futures in Education 11 564-74
[9] Olitsky S 2006 Cultural Studies of Science Education 1 745-66
[10] Hull M and Uematsu H 2018 GIREP-MPTL
[11] 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
[12] Perceived Choice and Awareness of Self Scale (PCASS) selfdeterminationtheory.org/pcass/
[13] Adams WK et al. 2006 Physical Review Special Topics-Physics Education Research 2 010101
[14] Wilson M 2004 “Constructing measures: An item response modeling approach”
[15] PRINCIPAL COMPONENT ANALYSIS support.sas.com/publishing/pubcat/chaps/55129.pdf
[16] Kaiser HF 1960 Educational and Psychological Measurement 20 141-151
[17] Lave J and Wenger E 1991 “Situated learning: Legitimate peripheral participation”
[18] Greeno JG 1997 Educational Researcher 26 5-17
[19] Peters GJ 2014 European Health Psychologist 16 56-69