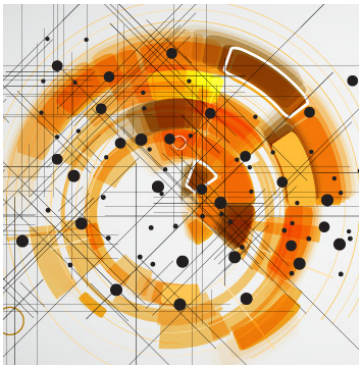



## Abstract

Traditional classroom assessment practice often leaves students out of the conversation, exacerbating the unequal power distribution in the classroom. Viewing classrooms as autonomy-inhibiting is known to influence students' psychosocial wellbeing as well as their academic achievement. This is especially relevant in STEM fields where marginalized populations experience disproportionate rates of attrition and success. The current paper describes the use of a simple, piloted classroom intervention with the intention of incorporating student voice in the co-creation of assessment criteria, relative to participation evaluation. Undergraduate students in a STEM research methods design course took part in the intervention. Pre- and post-tests were administered to understand what effects, if any, the intervention had on student perceptions of power and attitudes towards assessment. Perceptions of power significantly increased from pre- to post-test, and qualitative feedback from the intervention were overwhelmingly positive. Limitations and suggestions for future research are also discussed.



## AUTHORS

Manisha Kaur Chase, M.A.
*University of California Los Angeles*


# Student Voice in STEM Classroom Assessment Practice: A Pilot Intervention

Traditional classroom assessment practice is rarely known to involve students in the conception of assessment purpose or design (Falchikov, 2004). It follows that such an exclusion of student voice in assessment exacerbates the unequal power dynamics students experience in the classroom (McCroskey & Richmond, 1983; Sidky, 2017). This is especially important in Science, Technology, Engineering and Math (STEM) fields where marginalized populations experience disproportionate rates of attrition and greater barriers to success. To address these dynamics and to afford students more agency in their learning, recent higher education trends, particularly in STEM fields, show a shift towards student-centered pedagogy including the flipped classroom model and inquiry-based instruction, among others. Researchers argue, however, that assessment has been "neglected" in the pursuit of these strategies (Wanner & Palmer, 2015). The current paper describes the use of a simple, piloted classroom intervention with the intention of incorporating student voice in the co-creation of assessment criteria, relative to participation evaluation, and the larger goal of addressing some of the power imbalances and inequities present.


## CORRESPONDENCE

***Email***
mkc012@g.ucla.edu


### Student Voice & Co-Creation

"Student voice" is referred to here as "efforts that strive to redefine the role of students in educational research and reform" (Cook-Sather, 2006). The term was originally defined in the context of K-12 education but has also been applied to higher education (Brooman et al., 2015; Monsen & Cook, 2017). While attempts to involve students in

educational decisions at the university level is not a recent phenomenon, there has been as of late, a push to differentiate the types and respective consequences of current methods. For example, Bovill and Woolmer (2019) made a distinction between instructor and student "co-creation *of* the curriculum" versus "co-creation *in* the curriculum." Co-creation of the curriculum refers to that which occurs prior to the implementation of practice and often only includes fewer student voices in the process. In contrast, co-creation in the curriculum occurs during implementation of practice and typically engages a whole classroom of student voices. The latter strategy is least-often used in research and practice in higher education for various logistical challenges such as classroom size and time constraints. This is evident in the assessment realm, too, as studies that attempted to bring students into the assessment conversation revealed decisions about assessment practice having been made in isolation or prior to the beginning of a course (Pretorius et al., 2017; Wanner & Palmer, 2015).

**The aim of incorporating student voice into classroom matters—specifically that of assessment—is to trouble the notion of power norms and bring marginalized voices to the fore.**

The aim of incorporating student voice into classroom matters—specifically that of assessment—is to trouble the notion of power norms and bring marginalized voices to the fore; as demonstrated by previous work that specifically aimed to involve students in assessment practice (Bovill, 2020; Guberman, 2020; Deeley & Brown, 2014). However, in the case of co-creation of the curriculum, it is only the select students (typically those who have already taken a course, for example) that are privileged with the opportunity to have their voice heard as a result of their retrospective experience. Thus, while one could argue that student voice is being considered, the voices of those students *currently* in a given course are not consulted, which does little to challenge the power dynamic from their perspective. In this way, there is a need for co-creation in the curriculum that actively seeks to acknowledge student voice as their experiences in the classroom take place in order to provide enhanced inclusion and a more equitable learning environment. In doing so, such an approach may also serve to mitigate the power disparity students perceive in the classroom.

## STEM Context

The current study is not only concerned with the intersection of student voice and power in classroom assessment practice, but more specifically, in the STEM classroom. STEM disciplines are known to be "cutthroat" in nature (McGee, 2016) with "individualistic weed-out culture" (Daily et al., 2007) that are largely grounded in white, middle-class, masculine norms (Fabert et al., 2011). Retention in STEM is low across the board but known to disparately affect ethnic and gender minoritized groups (Chen & Soldner, 2013).

At the heart of many explanatory factors that attempt to narrow down the origins of this phenomenon (including a lack of belongingness, loneliness, and powerlessness) are assessment practices and grades which not only neglect to include these marginalized voices in their conception, but also expose the "gatekeeping" function of these assessments and the larger gatekeeping function of STEM as a whole (Committee, 2016). This perceived distance between students' respective identities and STEM content is reflected in assessment practices that are not only traditional in nature (quizzes and tests relying heavily on memorization techniques), but also "devoid of deep connection" to real life given their focus on isolated facts (Martin-Hansen, 2018; Momsen et al., 2010). In other words, despite the increasing diversity of undergraduate student populations, STEM classrooms often operate on implicit curriculum that reinforce the perceived objectivity of the disciplines themselves. Therefore, STEM is a context that may especially benefit from considering the need for student voice and co-creation in classroom assessment, but it is also one that requires a unique consideration of its culture, including hegemonic norms and the way those norms are perceived and acted upon by both instructors and students. The current paper and pilot intervention answers the call for inviting student voice and co-creation in the STEM classroom towards developing student autonomy (Bovill & Woolmer, 2019; Evans & Boucher, 2015), gauging psychosocial aspects of a classroom from the student perspective (Drewes et al., 2019), and engaging in assessment that explicitly acknowledges power stratification in the classroom and seeks to empower students (Bain, 2010).

## Theoretical Framework

Self-determination theory (SDT) is used here to highlight the need for examining autonomy development in the assessment realm. SDT is a theory of motivation which dichotomizes motivation as autonomous versus controlled (Deci & Ryan, 2008; Deci & Ryan, 1985). Autonomous (or intrinsic) motivation occurs when an individual experiences volition, through internal factors free from outside pressure or reward, versus controlled (or extrinsic) motivation in which an individual's actions are a result of external rewards or punishments.

In education, SDT argues that student motivation to learn can be explained by how well teachers encourage individual growth– including via autonomy. When an educational environment is perceived as having no room for control, self-determination—and consequently, motivation—experience decline (Deci et al., 1989). With a focus on choice and control, SDT situates the way a particular context enables or inhibits student voice/co-creation and its subsequent effects on individual autonomy (Ryan & Niemiec, 2009).

Student autonomy relative to assessment would require students having the opportunity to "take charge" (Holec, 1981) of those assessments that contribute to and evaluate their learning. In fact, an SDT approach to assessment would "actively empower and support change from within" (Ryan & Weinstein, 2009). However, the current state of assessment affairs—particularly those summative in nature—are predominantly conceived and implemented by instructors. Without experience in the assessment realm, students are left with a hierarchy that perpetuates student dependency on teacher judgement (e.g., grades) which may stifle student autonomy and exaggerate the power stratification (Sadler, 1989). However, by actively incorporating student voice in the classroom and increasing perception of choice, student autonomy development may flourish. Greater sense of student autonomy has been linked with greater internalized motivation, as well as better academic outcomes (Black & Deci, 2000; Chirkov & Ryan, 2001). In order to reap the benefits of empowering students, it becomes necessary not only to understand but also to address the power dynamics in a classroom—including how the dynamics are perceived, produced, and sustained—in attempts to engage students in the learning process in ways that bolster their motivation and sense of autonomy.

## Current Practice

**In other words, despite the increasing diversity of undergraduate student populations, STEM classrooms often operate on implicit curriculum that reinforce the perceived objectivity of the disciplines themselves.**

Current research demonstrates a variety of ways students have been involved in flexible, student-centered assessment practices. For example, a study by Pacharn et al. (2013) examined the effects of allowing undergraduate students the flexibility to allocate weights to the various assessments in their course. The full-flexible group (allowed to adjust weights until the end of the semester) reported significant increases in motivation, attitudes towards motivation, and higher academic outcomes than their early-flexible (only allowed to choose their weights at the beginning of the semester) and control group peers. In another example of flexible assessment, undergraduate students were given the opportunity to choose between multiple grading schemes for their course (Rideout, 2018). While choice of scheme had no significant correlation with final course outcomes, 79% of students reported being satisfied with having the choice.

These studies do provide an element of student choice in the assessment realm but have faltered in acknowledging student voice and supporting true student autonomy in the following ways. Firstly, assessment interventions that allow students to alter grading schemes and weightings serve only to reinforce grades and suggest that the conversation about assessment be limited to these quantifiers, rather than about the learning process as a whole. Moreover, in other examples where students were given the opportunity to choose a topic within an assignment or project (Bullen, 2012; Vandiver & Walsh, 2010), or provide feedback on a rubric (Fletcher & Shaw, 2012) they did so within the confines of an existing assessment format. In other words, students were not involved in the discourse on what purposes the assessments would serve, or which assessments might best serve said purposes. Students were involved only *after* these decisions had been determined. Finally, it should be noted that studies in this area rarely probe students' perceptions of power as a result of being involved in assessment dialogue. In one example of a partnership approach

to assessment, researchers found students expressed increased motivation and engagement as a result of being part of a more "democratic" assessment process (Deeley & Bovill, 2017). These data, however, were collected from qualitative student open-ended responses, and were not prefaced with a baseline measure of students' perceptions prior to being engaged in the process.

## Participation Evaluation

Engaging student voice in assessment practice may be an intimidating and confusing activity for students (Hewitt-Taylor, 2001; Monsen et al., 2017). Additionally, while dated, there are concerns that students do not have sufficient content knowledge in order to be involved in assessment practice (Falchikov, 2004; French et al., 1959). Thus, the current study suggests involving student voice in an area of evaluation that does not require content mastery: classroom participation.

In a survey of one urban university, 93% of all courses included participation as part of overall course grades (Bean & Peterson, 1998). A survey of 520 instructors at another large, state university revealed 82% of faculty reportedly including participation in their syllabus, with only 25% of these professors actually providing criteria and grades for participation (Rogers, 2013). Why were grades not formally assigned to participation? One professor noted, "'I believe that different students learn in different ways, and forcing quiet students to talk in class is obnoxious and likely to be counterproductive, e.g., superficial participation just for points'" (p. 18). While on the surface this appears to be a considerate reflection from the instructor's point of view, their comment serves to ignore the student voice regarding what good participation may look like (i.e., in the case of the "quiet student"), and rests on an esoteric, normative assumption that good participation is represented solely by speaking up in class. By dismissing the evaluation altogether, the professor missed an opportunity to engage student voice through co-creation in assessment, in order to develop criteria that may be representative of the diversity of students and their respective needs and preferences. Such a critical perspective—that criteria must take into account minoritized students, cultural conflicts, and issues of representation and power in the classroom, relative to participation evaluation—has also been cited in the literature (Meyer & Hunt, 2011; White, 2011).

**[This] study set out to explore how an intervention that highlights undergraduate STEM students' voices through co-creation in participation criteria might affect student perceptions of autonomy support (power) and attitudes towards assessment.**

Given the prevalence of participation evaluation in course grades and the discord between its mention and its actual evaluation, I argue this is an appropriate area to begin incorporating student voice in assessment practice.

While research has shown previous attempts to include students in the creation of participation criteria, these attempts have failed to have students come up with the operationalization of participation skills and grading, and have not yet shown the impact on student perceptions of psychosocial outcomes (Dancer & Kamvounias, 2005). Thus, the current pilot study set out to explore how an intervention that highlights undergraduate STEM students' voices through co-creation in participation criteria might affect student perceptions of autonomy support (power) and attitudes towards assessment.

# Method

This quasi-experimental pilot study employed a pretest-posttest design to uncover the potential effects of an assessment intervention on student perceptions of their classroom. An informed consent waiver was distributed to all students outlining their participation in the study. IRB approval was obtained in order to disguise the intervention's true purpose and risk biasing survey responses wherein students were told the study was vaguely aimed at understanding student perceptions of their classroom. Following the study, students were debriefed as to the true intentions of the study, including explicitly addressing the intervention and its relation to perceptions of power and attitudes towards assessment specifically.

## Participants and Context

The pilot intervention was carried out with a total of 21 undergraduate students (ages 18-22) from an upper-division, STEM, research methods design lab course at a large,

public university. The sample consisted of seven self-identified Caucasian/White students, four each self-identifying as Latinx/Hispanic, East Asian, and Multiethnic, respectively, while the remaining students self-identified as Southeast Asian or Middle Eastern. Three-quarters of the sample self-identified as female while the remaining self-identified as male. Three students identified as being transfers, two reported international student status, and seven identified as first-generation college students. Only one student reported having a learning accommodation. Overall, these demographics are reflective of the specific STEM major at this institution. The course itself was comprised of a weekly asynchronous lecture taught by the instructor of record while the lab course met synchronously with graduate Teaching Assistants (TA)—twice weekly for two hours each over the course of an academic quarter (10-week period). Given the ongoing global COVID-19 pandemic at the time of this course, the usual in-person curriculum was adapted for online instruction (including the synchronous lab discussions).

Course grades were a compilation of the lecture exams (40%) and the lab assignments and activities (60%). Participation consisted of 20% of students' lab grades, which included a day-to-day evaluation from the TA in addition to an end of quarter peer-assessment relative to a group project implemented in the course. The current intervention focused on the criteria which informed the day-to-day evaluation portion of students' participation evaluation. Typically, the criteria for what constitutes good participation is not mandated from the instructional team and thus, left up to the discretion of the TA. In this pilot study, the researcher capitalized on this flexibility in an attempt to bring students into the assessment conversation.

## Measures

Power was operationalized as students' perceptions of autonomy support from their instructor in addition to their perception of having voice in the classroom. The 6-item "Learning Climate Questionnaire" [(LCQ); Williams & Deci, 1996] was adapted for the purpose of this study [$\alpha$ =.65, $p$=.001; acceptable reliability as compared to $\alpha \geq .60$ (van Griethuijsen, 2015)]. Participants were prompted to "think about the way you are assessed by your TA and respond to the following prompts in regards to that assessment experience." Sample items included: "I feel that my TA provides me choices and options" and "My TA conveyed confidence in my ability to develop assessment criteria." Responses fell on a Likert scale from 1-7 with 1 representing "Strongly Disagree" and 7 representing "Strongly Agree." Item responses were aggregated into a single perception of power score for each participant, where higher aggregated scores suggested increased perceived autonomy support/power.

**A final purpose of the intervention was to provide students a holistic experience—from the very beginning of determination of purpose to the 'end result' of grading itself—of assessment in the classroom.**

Meanwhile, student attitudes toward assessment was operationalized as students' preference and beliefs regarding assessment in their classroom. A 5-item version adapted from the "Attitudes towards Grading System" scale (Pacharn et al., 2013) was used to gauge student attitudes ($\alpha$ =.24, $p$ >.05). Sample items included: "I liked how the grading scheme employed in this course, with respect to participation, was determined" and "I believe that allowing students to participate in designing the grading scheme (e.g., in relation to participation) in a course wastes students' time that could be better spent working on the course material." Participants responded on a 7-point Likert scale with 1 indicating "Strongly Disagree" and 7 indicating "Strongly Agree." Once again, item responses were aggregated into a single attitude towards assessment score for each participant.

Finally, in order to understand the qualitative experience of the intervention, a survey administered half-way through the quarter consisted of a free response where students had the opportunity to describe how the experience of being involved in assessment development made them feel, what effect it had on their perceptions of the classroom/instructor, what they enjoyed about the experience, and what might be used to improve the intervention. This provided anecdotal data on students' experience of and suggestions to improve the intervention.

## Procedure

The overall aim of the pilot intervention—as outlined in detail below—was to involve student voice in classroom assessment practice. More specifically, the intervention sought

to achieve the following: meaningfully engage student voice in the assessment development process through the co-creation of participation evaluation criteria and provide students an opportunity to stray away from the historical dependence on instructors for assessment evaluation. Additionally, as a result of having to create the criteria in addition to apply it via self-assessment, a final purpose of the intervention was to provide students a *holistic* experience—from the very beginning of determination of purpose to the 'end result' of grading itself—of assessment in the classroom.

The pre-survey was administered on the first day of lab during Week 1, followed by the qualitative survey gauging the intervention process at Week 5, and finally, the post-survey at Week 10.

## The Intervention

The proposed intervention took place on the first day of class as the Teaching Assistant went over the syllabus policies of the lab. Following the pre-survey, the TA prefaced the intervention by discussing the challenges of assessment in school contexts. More specifically, the challenge of attempting to measure something, unlike weight or height, that is not tangible. The TA cited a history of researchers, policy makers, professors, etc. working to hone assessment practices to make them fair, valid for their outset purpose, and reliable. However, the TA noted this iterative process has often failed to incorporate student voice in what is classified as important. The TA then expressed that in the current context, they wanted to give that opportunity to students such that they may co-create meaning of one aspect of assessment in the course: participation.

**The final five skills included: written communication, verbal communication, engagement, critical thinking, and teamwork.**

In guiding students to think about assessment purposes at large, the TA first asked the class why participation may be a part of their grade (when it seemingly has nothing to do with STEM content). Probing questions included: "What might participation (in all its forms) be representative of? What skills might we be assessing when it comes to the various aspects of assessments?" Essentially, the TA asked students to think relative to *this specific course* what skills students valued and wanted to foster in their space. Using the web-based response site, *Mentimeter*, students were asked to record three larger skills they believe participation represents. These were then generated into one cohesive word cloud (see Figure 1). Using the word cloud, the TA engaged students in a discussion of consolidating these into a handful of larger skills participation would be representative of in the course. The final five skills (read here as purposes of participation evaluation in this context) included: written communication, verbal communication, engagement, critical thinking, and teamwork.

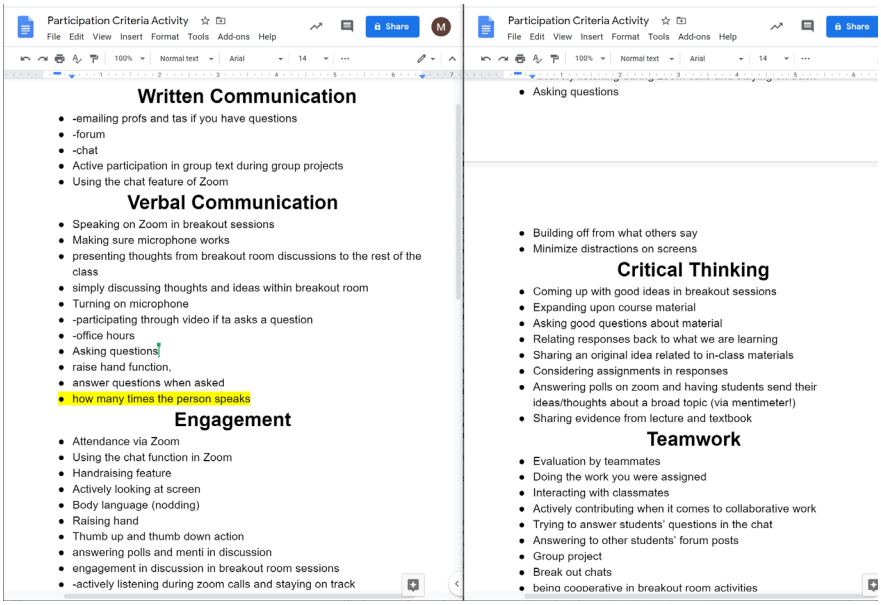
Once these skills were established, the TA split students up into smaller groups to discuss what concrete behaviors might be representative of each of those skills (i.e., a skill of being respectful in the classroom might have a concrete behavior of not being distracted via cell phone or laptop use). Each group elected a scribe who transferred the group notes into a shared Google Doc (see Figure 2). The TA then had students return from their groups and take a few minutes to review the criteria their peers had constructed. This was followed by a facilitation of how students felt about the criteria in general, and if there were any criterion students would like to make more specific, or perhaps, remove.

In this iteration of the intervention a student did in fact raise a concern. Under the skill "verbal communication," one group had suggested, "how many times a student speaks," as a potential criterion to be used. The student contested that perhaps, particularly in an online format, asking students to speak may be uncomfortable for some and that additionally, quantifying verbal communication might lead to students "speaking for the sake of it" while not lending any meaningful contribution to discussion. The TA then asked if there were any counterpoints to the concern raised, and an anonymous poll was employed in which the students unanimously voted to expel that criterion.

Following this process, the TA created a polished version of the criteria and noted that this is what would be used to evaluate student participation in lab (see Figure 3). In order to scaffold the assessment development experience for students, a scale for grading was suggested for the criteria created. The TA explained how the criteria would align with the three

Figure 1
*Sample Student Word Cloud of Meta-skills Represented by Participation*


Figure 2
*Sample Working Google Document of Purposes and Criteria*




participation points students could earn for each lab session (a combination of the quantity and quality of student engagement). To avoid confusion, it was cautioned that students were not expected to engage in *all* criteria. Overall, the process took approximately 20 minutes. This concluded the first part of the intervention.

The second part of the intervention took place during Week 5 of the quarter where students were reminded of the criteria and had an opportunity to engage in self-assessment. First, students were administered the qualitative survey which probed affective feelings about the intervention in addition to feedback on the process. Thereafter, students were asked to qualitatively self-assess how they feel they had lived up to the participation criteria they developed. Finally, space was provided for students to indicate any criterion they would like to change or add given their experience in the first half of the course. Following class, the students were notified by the TA of the score they had accumulated in the first half of the quarter according to the TA's evaluation using the criteria. This provided one way in which students could understand how the criteria they developed had resulted in their actual participation grade as well as an opportunity to show that the criteria they created were in fact

Figure 3
*Sample Finalized Criteria*

Participation Criteria for [redacted]

| Written Communication | Verbal Communication | Engagement | Critical Thinking | Teamwork |
|---|---|---|---|---|
| • Emailing the professor or TAs if you have questions<br>• Using the CCLE forum to ask questions/pose answers<br>• Using the chat function on Zoom<br>• Active participation in group text during group projects | • Using the raise hand function or simply raising hand<br>• Sharing thoughts, ideas, and questions verbally in Zoom breakout sessions/main room<br>• Making sure microphone works<br>• Attending office hours<br>• Providing answers to questions posed by TA and classmates | • Attendance via Zoom<br>• Using the chat function, thumbs up, or raising hand in Zoom<br>• Using body language such as actively looking at screen, nodding, thumbs up, or raising hand<br>• Answering polls and Mentimeter questions<br>• Actively listening<br>• Asking questions<br>• Building off of what the TA or peers say<br>• Minimize distractions on screens | • Coming up with good ideas in breakout sessions/main room<br>• Expanding upon course material, including asking questions<br>• Relating responses back to what we are learning in lecture<br>• Sharing an original idea related to in-class materials<br>• Considering assignments in responses<br>• Sharing evidence from lecture and textbook (and potentially, real life) | • Evaluation by teammates<br>• Doing the work/actively contributing to what you were assigned by your group<br>• Interacting with classmates during breakout sessions<br>• Staying on task within groups<br>• Trying to answer peers' questions in the chat and/or forum posts |

being used by the TA. Sixty-two percent of the class at this time reported being satisfied with the criteria, while the remaining students had comments relative to the quantity of criteria. In the words of one student: it was "hard to narrow down the main points." The TA addressed these concerns citing the need for inclusivity such that the breadth of criteria was meant to cater to individual strengths in participating, rather than being a punitive system that required students to participate using all suggested criteria. Comments and suggestions were welcomed but appeared to have been satisfied with the aforementioned explanation. Finally, students were given the post-survey in Week 10.

## Results

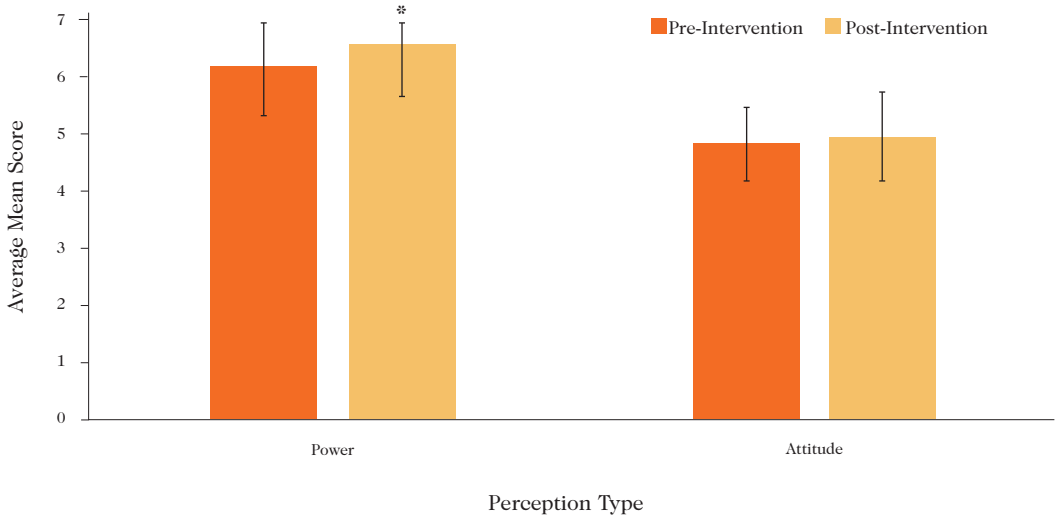
Figure 4 presents the average aggregate score of perception of power and attitudes towards assessment pre- and post-intervention, respectively. The average perception of power during the pre-test ($M$=6.02, $SD$=0.79) was lower than the average reported in the post-test ($M$=6.44, $SD$ =0.88). Similarly, the average attitude towards assessment was lower in the pre-test ($M$=4.74, $SD$ =0.65) than that of the post-test ($M$=4.88, $SD$ =0.79). A Shapiro-Wilk test for normality revealed a non-normal distribution of both pre- and post-test power scores ($p$=.005; $p$=.000) and a normal distribution for both pre- and post-test attitude scores, respectively ($p$=.38; $p$=.19).

**The average perception of power during the pre-test was lower than the average reported in the post-test.**

In order to compare within-subject differences, a related-samples Wilcoxon Signed Rank Test for nonparametric data was conducted for power scores, and a paired-samples t-test for attitudes towards assessment were conducted to reveal any significant changes in mean perceptions reported from the pre-post surveys. These tests revealed a significant difference in perceptions of power from pre-survey to post-survey ($Z$=91, p=.015), where student perceptions of their own power in the classroom were higher, or closer to "Agree," on the post-test compared to the pre-test. The effect size for this analysis ($d$ =-1.12) was found to exceed Cohen's (1998) assumption for a large effect ($d$ =0.80). Finally, no significant difference for attitudes toward assessment were found from pre- to post-survey ($t(20)$ =-.69, $p$=.50; $d$ =-0.15).

The open-ended responses were analyzed using an inductive, open coding process in order to allow similar themes to arrive from data itself, rather than employing a pre-conceived, deductive coding scheme to student experience. The first round focused on those experiences that students appeared to share during the process, the second round consisted of consolidating those themes, and the final round searched for exceptions to the themes.

Figure 4
*Perception of Power and Attitudes Toward Assessment Scores Pre and Post Intervention*




When asked about how students felt they had lived up to the criteria, many responded they felt they had adequately met a fair number of criteria—particularly those criteria that played to their strengths. One student felt they had not met the criteria, citing difficulty focusing given the switch to virtual instruction. In contrast, another student cited they felt they were able to meet the criteria as the process provided "leeway during this time of uncertainty." This comment was echoed in a subsequent response to which a student said they enjoyed that their voice was included in the process "especially [given the] uncertain time."

Relative to affective reactions to the intervention, students had mixed reviews. One, for example, said the process made them feel "neutral" and didn't feel "it made much of a difference" to their classroom experience. Another cited the experience as "a little uncomfortable." Overwhelmingly though, students felt positive about the process. One student said the process made them feel "Good!," while another appreciated "the instructors taking the time to allow us to build our own criteria and see what is working." That student went on to say: "it shows a lot of respect for our time and opinion."

**One student posed that the process: "helped me understand the why behind some of the criteria, for example sometimes we are graded on things that don't seem important but being able to create what we are graded on makes sense."**

When asked how the process affected student perceptions of the classroom and/or the instructor, students alleged feeling "listened to," which prompted a more positive outlook of the course because, "it showed that the instructor wanted to get our input." This led to perceptions of "really respecting" the instructor via students' perception that "they [the instructor] actually care." In fact, one student claimed they felt "closer to [their] TA than most of [their] TA's in the past." Finally, one student posed that the process: "helped me understand the why behind some of the criteria, for example sometimes we are graded on things that don't seem important but being able to create what we are graded on makes sense."

Finally, students responded to questions about what worked about the process and what could be improved. The majority of students felt the process worked as is, and particularly enjoyed breaking up into smaller groups in order to come up with criteria. This echoes findings that students appreciate the informality of small groups when wrestling with new ideas (Monsen et al., 2017). As mentioned earlier, however, a couple of students noted the extensiveness of the criteria as potentially overwhelming, with a lack of clarity as to which specific criterion to prioritize. This is consistent with research on classroom choice which cites marginal utility relative to the amount of choice offered (Patall et al., 2008). As a process, one student remarked, quite honestly, it was a "reminder to stay involved [but] feels a little like nagging though." With respect to improving the intervention, one student asked for a clearer understanding of how criteria mapped on to the grading scale, while another said they could have done with a shorter intervention on the first day of class.

## Discussion

The current pilot study explored the effects of an intervention that sought to engage student voice in assessment (relative to participation evaluation) on their perceptions of power and attitudes towards assessment in a STEM classroom. Students reported a significant difference in their perceptions of power prior to and after implementation of the intervention, such that students felt stronger autonomy support in the classroom following the intervention. This follows findings in the realm of co-creation in curriculum where students express increased autonomy and motivation as a result of the process (Bergmark & Westman, 2015; Deeley & Bovill, 2017). A limitation to these findings includes an absence of control group which makes it difficult to isolate the intervention as the sole reason for differences found in student perceptions of power. Moreover, as students took the pre-survey on the first day of instruction their experience and ability to comment on that experience in the classroom, was limited. Nonetheless, these findings are encouraging considering the significant increase in student perception of their own power given the statistically small sample size and mere 10-week study period.

**Overall, the intervention process appeared to bolster students' general perceptions of the classroom given their comments of feeling "listened to," "cared" about, and even, feeling "closer" to the instructor and the classroom.**

In terms of the intervention experience, students largely appeared to find the process positive and helpful. While not a focus of the study, many students addressed the "uncertain times" that accompanied online instruction amidst a pandemic and cited the intervention as particularly useful therein. This has implications for virtual instructional methods, ways students can be involved in assessment dialogue despite not having a traditional learning environment, and how doing so may positively affect classroom perceptions.

This experience, however, was not intuitive for all participants. The discomfort noted from one student is perhaps a reminder of the traditional classroom perceptions of power where teachers are considered "the sole authority" in the classroom, and students "surrender" to that power as part of an "unwritten contract" (Sidky, 2017). By asking students to participate in the assessment realm, one they are not historically a part of, may lead to feelings of being unprepared and uncomfortable. This reinforces the idea of needing to engage student voice in assessment co-creation such that students become comfortable, and are ultimately prepared for a "lifetime of assessing their own learning" beyond the classroom (Boud & Falchikov, 2006, p. 400). The current intervention appears to begin this process, as cited by the student who said they now "understand the *why* behind criteria," as opposed to their previous experiences with grading and assessment. This comment is echoed in existing literature which suggests that involving students in the assessment process can provide an experience that "totally deepened my [student participant's] learning," by providing a metacognitive lens for what happens in the classroom and how these processes affect students' own learning (Cook-Sather, 2018, p. 927). While this intervention did not cover content-based assessment, it appeared to provide a small stepping-stone in getting students to understand the 'behind-the-scenes' of one aspect of assessment practice. Future iterations of this intervention should look to expand such work toward content-based assessments in the classroom.

Overall, the intervention process appeared to bolster students' general perceptions of the classroom given their comments of feeling "listened to," "cared" about, and even, feeling "closer" to the instructor and the classroom than they perhaps otherwise would have been (despite the virtual instructional setting). These findings of dialogue being advantageous towards students' perception of power in the classroom follow McLean's (2018) suggestion of explicit dialogue being essential to negotiate power and build trust between students and faculty. More generally, co-creation and partnership approaches in the classroom—validating and encouraging student voice and respective experience—begin the work of highlighting and subsequently tackling systems that lead to inequity (Cook-Sather, 2019). In fact, co-creation and the dialogue that accompanies it, has additionally been found to bolster the confidence of instructors towards explicitly addressing classroom inequities. Gauging faculty perceptions in future research of the co-creation process may be useful in validating such findings.

Future work using this intervention is currently underway to understand if the effects found in this pilot are replicable in traditional, in-person STEM classrooms, with larger sample sizes, and over longer periods of instructional time. More research in this area is necessary to understand how such an intervention affects perceptions across student identities, particularly

historically marginalized populations in higher education and STEM fields where the voices of such populations are not often represented in both the content and assessment of such content. Finally, future research should measure the effects of this intervention on other relevant outcome variables such as academic performance, feelings of belongingness, self-regulation, assessment anxiety, and retention/success in STEM.

## Conclusion

The current pilot study aimed at engaging student voice in assessment practice, with the explicit purpose of studying student perceptions of power and attitudes towards assessment. While this study begins with involving students in non-content related assessment practice in a STEM classroom, the hope is to create an empirical foundation upon which research and practice can meaningfully incorporate student voice in content-related designs as well. More work in this area may help develop and validate this simple intervention for faculty to engage students in assessment without a complete overhaul of their existing assessment practices. Partnering with students at the assessment table may serve to empower and improve perceptions of the classroom, toward the end of fostering a more equitable learning environment for all students.

## References


Bain, J. (2010). Integrating student voice: Assessment for empowerment. *Practitioner Research in Higher Education, 4*(1), 14-29.

Bean, J. C., & Peterson, D. (1998). Grading classroom participation. *New Directions for Teaching and Learning*, 1998(74), 33-40. https://doi.org/10.1002/tl.7403

Bergmark, U., & Westman, S. (2016). Co-creating curriculum in higher education: Promoting democratic values and a multidimensional view on learning. *International Journal for Academic Development, 21*(1), 28-40.

Black, A. E., & Deci, E. (2000). The effects of instructors' autonomy support and students' autonomous motivation on learning organic chemistry: A Self-Determination Theory perspective. *Science Education, 84*(6), 740-756.

Boud, D., & Falchikov, N. (2006). Aligning assessment with long-term learning. *Assessment & Evaluation in Higher Education, 31*(4), 399-413.

Bovill, C. (2020). Co-creation in learning and teaching: The case for a whole-class approach in higher education. *Higher Education, 79*(6), 1023-1037.

Brooman, S., Darwent, S., & Pimor, A. (2015). The student voice in higher education curriculum design: Is there value in listening? *Innovations in Education and Teaching International, 52*(6), 663-674. 910128

Bullen, H. A. (2012). Student-driven independent research projects: Developing a framework for success in analytical chemistry. *Analytical and Bioanalytical Chemistry, 404*(4), 927-930.

Chen, X., & Soldner, M. (2013). *STEM Attrition: College students' paths into and out of STEM fields* (p. 104). Washington, DC: National Center for Education Statistics.

Chirkov, V. I., & Ryan, R. M. (2001). Parent and Teacher Autonomy-Support in Russian and U.S. Adolescents: Common Effects on Well-Being and Academic Motivation. *Journal of Cross-Cultural Psychology, 32*(5), 618-635.

Cohen, J. (1988). *Statistical power analysis for the behavioral sciences* (2nd ed). Hillsdale, N.J: L. Erlbaum Associates.

Committee on Barriers and Opportunities in Completing 2-Year and 4-Year STEM Degrees, Board on Science Education, Policy and Global Affairs, National Academy of Engineering, & National Academies of Sciences, Engineering, and Medicine. (2016). *Barriers and Opportunities for 2-Year and 4-Year STEM Degrees: Systemic Change to Support Students' Diverse Pathways*. (S. Malcom & M. Feder, Eds.). Washington, D.C.: National Academies

Cook-Sather, A. (2006). Sound, Presence, and Power: "Student Voice" in Educational Research and Reform. *Curriculum Inquiry, 36*(4), 359-390.

Cook-Sather, A. (2018). Listening to equity-seeking perspectives: how students' experiences of pedagogical partnership can inform wider discussions of student success. *Higher Education Research & Development, 37*(5), 923-936.

Cook-Sather, A. (2019). Respecting voices: How the co-creation of teaching and learning can support academic staff, underrepresented students, and equitable practices. *Higher Education*, 1-17.

Daily, S., Eugene, W., & Prewitt, D. (2007). The development of social capital in engineering education to improve student retention. In Proceedings of the ASEE Southeast Section Conference. Louisville, Kentucky.

Dancer, D., & Kamvounias, P. (2005). Student involvement in assessment: A project designed to assess class participation fairly and reliably. *Assessment & Evaluation in Higher Education, 30*(4), 445-454.

Deci, E., Connell, J. P., & Ryan, R. (1989). Self-determination in a work organization. *Journal of Applied Psychology, 74*(4).

Deci, E. L., & Ryan, R. M. (2008). Self-determination theory: A macrotheory of human motivation, development, and health. *Canadian Psychology/Psychologie Canadienne, 49*(3), 182-185.

Deci, E., & Ryan, R. (1985). *Intrinsic motivation and self-determination in human behavior*. Plenum.

Deeley, S. J., & Bovill, C. (2017). Staff student partnership in assessment: enhancing assessment literacy through democratic practices. *Assessment & Evaluation in Higher Education, 42*(3), 463-477.

Deeley, S. J., & Brown, R. A. (2014). Learning through partnership in assessment. *Teaching and learning together in higher education*, *1*(13), 3.

Drewes, A., Scantlebury, K., & Soslau, E. (2019). Listening to the Missing Voices: Students' Perspectives on Coteachingdf. *Research & Practice in Assessment, 14*(Winter), 5-18.

Evans, M., & Boucher, A. R. (2015). Optimizing the Power of Choice: Supporting Student Autonomy to Foster Motivation and Engagement in Learning. *Mind, Brain, and Education, 9*(2), 87-91.

Fabert, N., Cabay, M., Rivers, M. B., Smith, M. L., & Bernstein, B. L. (2011). Exaggerating the typical and stereotyping the differences: Isolation experienced by women in STEM doctoral programs. *American Society for Engineering Education, 16*.

Falchikov, N. (2004). Involving students in assessment. *Psychology Learning and Teaching, 3*(2), 102-108.

Fletcher, A., & Shaw, G. (2012). How does student-directed assessment affect learning? Using assessment as a learning process. *International Journal of Multiple Research Approaches, 6*(3), 245-263.

French, R., Raven, B., & Cartwright, D. (1959). The bases of social power. *Classics of Organization Theory, 7*, 311-320.

Guberman, D. (2020). Grading together: Toward a partnership approach. *Teaching and Learning Together in Higher Education, 1*(29), 6.

Hewitt-Taylor, J. (2001). Self-directed learning: Views of teachers and students. *Journal of Advanced Nursing, 36*(4), 496-504.

Holec, H. (1981). *Autonomy and Foreign Language Learning*. Pergamon Press.

Martin-Hansen, L. (2018). Examining ways to meaningfully support students in STEM. *International Journal of STEM Education, 5*(1), 53.

McCroskey, J. C., & Richmond, V. P. (1983). Power in the classroom I: Teacher and student perceptions. *Communication Education, 32*(2), 175-184.

McGee, E. (2016). Devalued Black and Latino racial identities: A by-product of STEM college culture? *American Educational Research Journal, 53*(6), 1626-1662.

McLean, H. (2018). This is the way to teach: Insights from academics and students about assessment that supports learning. *Assessment & Evaluation in Higher Education, 43*(8), 1228-1240.

Meyer, K. R., & Hunt, S. K. (2011). Rethinking evaluation strategies for student participation. *Basic Communication Course Annual, 23*, 93-126.

Momsen, J. L., Long, T. M., Wyse, S. A., & Ebert-May, D. (2010). Just the facts? Introductory undergraduate biology courses focus on low-level cognitive skills. *CBE—Life Sciences Education, 9*(4), 435-440.

Monsen, S., Cook, S., & Hannant, L. (2017). Students as partners in negotiated assessment in a teacher education course. *Teaching and Learning Together in Higher Education, 1*(21).

Pacharn, P., Bay, D., & Felton, S. (2013). The impact of a flexible assessment system on students' motivation, performance and attitude. *Accounting Education, 22*(2), 147-167.

Patall, E. A., Cooper, H., & Robinson, J. C. (2008). The effects of choice on intrinsic motivation and related outcomes: A meta-analysis of research findings. *Psychological Bulletin, 134*(2), 270-300.

Pretorius, L., van Mourik, G. P., & Barratt, C. (2017). Student choice and higher-order thinking: Using a novel flexible assessment regime combined with critical thinking activities to encourage the development of higher order thinking. *International Journal of Teaching and Learning in Higher Education, 29*(2), 389-401.

Rideout, C. A. (2018). Students' choices and achievement in large undergraduate classes using a novel flexible assessment approach. *Assessment & Evaluation in Higher Education, 43*(1), 68-78.

Rogers, S. L. (2013). Calling the question: Do college instructors actually grade participation? *College Teaching, 61*(1), 11-22.

Ryan, R. M., & Niemiec, C. P. (2009). Self-determination theory in schools of education: Can an empirically supported framework also be critical and liberating? *Theory and Research in Education, 7*(2), 263-272.

Ryan, R. M., & Weinstein, N. (2009). Undermining quality teaching and learning: A self-determination theory perspective on high-stakes testing. *Theory and Research in Education, 7*(2), 224-233.

Sadler, D. R. (1989). Formative assessment and the design of instructional systems. *Instructional Science, 18*(2), 119-144.

Sidky, G. (2017). The power game: Power dynamics between the teacher and the students in a graduate seminar. *English Language Teaching, 10*(5), 179.

Vandiver, D. M., & Walsh, J. A. (2010). Assessing autonomous learning in research methods courses: Implementing the student-driven research project. *Active Learning in Higher Education, 11*(1), 31-42.

van Griethuijsen, R. A., van Eijck, M. W., Haste, H., den Brok, P. J., Skinner, N. C., Mansour, N., ... & BouJaoude, S. (2015). Global patterns in students' views of science and interest in science. *Research in science education, 45*(4), 581-603.

Wanner, T., & Palmer, E. (2015). Personalising learning: Exploring student and teacher perceptions about flexible learning and assessment in a flipped university course. *Computers & Education, 88*, 354-369.

White, J. W. (2011). Resistance to classroom participation: Minority students, academic discourse, cultural conflicts, and issues of representation in whole class discussions. *Journal of Language, Identity & Education, 10*(4), 250-265.

Williams, G. C., & Deci, E. L. (1996). Internalization of biopsychosocial values by medical students: A test of Self-Determination Theory. *Journal of Personality and Social Psychology, 70*(4), 767-779.