

**Building Social Capital by Balancing Voices in School Governance:
Results from an RCT**

Kenneth A. Frank*, Kaitlin Torphy*, John Lane and Dirk Zuschlag
Michigan State University
*equal first authors

We acknowledge the contributions of Sarah Galey, Tony Bryk, Richard Ingersoll, Brian Kalt, Kristi Bowman, Barbara Markle, Donald Peurach, Michael Flanagan, and Nancy Knight.

Ken Frank

MSU Foundation Professor of Sociometrics

Measurement and Quantitative Methods

Counseling, Educational Psychology and Special Education

Room 462 Erickson Hall

Michigan State University

East Lansing, MI 48824-1034

phone: 517-355-9567 fax: 517-353-6393. kenfrank@msu.edu

Abstract

Social capital has been found to have positive effects on teachers' instructional practices and student outcomes. We propose that schools can build social capital through an explicit school governance framework called Balancing Voices that concerns decisions about implementing and evaluating reforms, and evaluating teachers and administrators. In an RCT of role play simulations, we find that those assigned to a Balancing Voices approach reported higher levels of procedural fairness and legitimacy of authority figures than those assigned to "business as usual." Accordingly, schools that more explicitly and formally balance the interests of different stakeholders in their decision making may be able to cultivate greater flows of social capital to improve instructional practices and student outcomes.

Social capital among educators has been shown to contribute to the implementation of innovations (e.g., Coburn et al., 2012; Frank, Zhao & Borman, 2004; Penuel et al., 2009) and a positive work environment that affects student achievement (Bryk & Schneider, 2002; Ronfeldt et al., 2015; Sun, Loeb and Grissom, 2017).¹ Recent attention has focused on how teachers identify relevant social capital by accessing advice or knowledge in their networks (Coburn and Russell, 2008; Spillane, Kim & Frank, 2012; Wilhelm et al., 2016). Yet, it is less clear what schools can do to cultivate social capital flows. In this study, we focus on how a school’s decision-making affects its ability to cultivate social capital flows. In particular, we present the Balancing Voices Framework (BVF) for school decision-making, an approach that we show through simulations leads to more perceived fairness of decisions and legitimacy of those who make decisions. In turn, perceived fairness and legitimacy can lead to teachers’ investment in their schools and one another, contributing to flows of social capital that improve instruction and student outcomes.

We assume that educators join a school seeking to educate students and contribute to the common good of the school (Frank, Kim & Belman, 2010; Ingersoll & Perda, 2008; Ingersoll & Strong, 2011). But educators may encounter major forces from outside the school that reduce their motivation and commitment to their schools. First, multiple reforms can pit educators against one another in competition for resources and status (e.g., Honig & Hatch, 2004; Penuel et al., 2009; Powell, Farrar & Cohen, 1985). Second, the evaluation of teachers based on only what occurs in their classrooms (e.g., value

¹ We use the definition of social capital as “the ability of actors to secure benefits by virtue of membership in social networks or other social structures” (Portes, 1998, page 6). See also Lin (1999).

added scores or classroom observations) incentivizes teachers' investments in only their classroom, and not the classrooms of their colleagues (Grissom et al., 2014; Harris, 2009; Springer et al., 2011). Third, market forces that attract teachers to other schools create turnover that can reduce incentives to invest in others within the school (Holme, et al., 2017; Liang & Akiba, 2017). Fourth, fragmented input from parents can create competing demands on schools that tear at the social fabric within (Bidwell Frank & Quiroz, 1997; Powell, Farrar & Cohen, 1985).

We believe that schools can mitigate the effects of external forces through the way they make fundamental decisions to adopt reforms, to evaluate and retain personnel, and to organize parental/community input. Our approach is to present a framework for explicitly balancing the voices of the key stakeholders in schools: teachers, community members, the district, and administrators. Balancing the voices of stakeholders allows schools to respond to external forces while maintaining internal coherence and commitment. Ultimately, improving school governance can contribute to student learning by contributing to coordination and knowledge sharing among educators.

Theory: Cultivating Social Capital

Our approach extends existing research concerning the cultivation of social capital in schools. In their seminal work *Trust in Schools* (2002), Bryk and Schneider argued that social capital builds over time when actors engage in frequent, positive interactions and when they share a common vision for the school. More specifically, according to Bryk

and Schneider social capital is enhanced when interactions among teachers, principals, students, and parents demonstrate competence, high regard and respect for others, and personal integrity, and when all actors rally around the common purpose that schools are primarily about improving the learning conditions for students.

Critically, Bryk and Schneider (2002) concluded “school decision making must work for both adults and students alike. The key task for school leadership involves getting the balance right” (page 136). We contend getting the decisions right depends as much on the formal approach to school governance as on the character and subjective choices of specific leaders. Moreover, as schools welcome new leaders over time, stable, effective school governance may ensure a consistent balance in decision making. In particular, school governance based on a set of pre-established and balanced rules helps schools make difficult and sometimes contentious decisions to balance the interests of different stakeholders. Furthermore, explicit, formal governance pre-empts the need for ad-hoc decisions that are often not transparent or clearly rooted in underlying principles (Hornig, Klasik, & Loeb, 2010; Spillane & Lee, 2014; Spillane, Camburn & Stitzel, 2007).

Teacher Commitment to the School and School Decision-Making

To articulate our theory of how school governance can improve schools and learning, we turn to the recent literature on organizational decision-making and employee commitment. An organization could cultivate commitment by allocating resources to its members. This is the theory behind distributive fairness (Folger & Konovsky, 1989, page

115). In schools, salary limits shift the allocation of desirable resources to those other than salary (Bidwell & Kasarda, 1980). For example, a principal can reward a high performing teacher by giving her discretion in next year's student assignments, or can punish a low performing teacher by continually modifying her teaching assignment. These types of allocations concern fixed resources, such as teachers' preferred classes. Therefore, for every resource allocation, some will win and others will lose. As a result, there is a limit to how much teacher commitment can be broadly cultivated through distributive fairness alone.

Partly in reaction to the limitations of distributive fairness, recent literature in organizational psychology has attended to procedural fairness -- "the perceived fairness of the *means used to determine* those amounts [of resources allocated]" (Folger and Konovsky, 1989, page 115, emphasis added). Critically, procedural fairness has been extensively linked to organizational citizenship behavior defined as "acts of cooperation, helpfulness, suggestions, gestures of goodwill, altruism, and other instances of what we might call citizenship behavior." (Folger & Konovsky, page 653; see also Smith, Organ & Near, 1983; Organ, 1988). Organizational citizenship behavior aligns closely with Bryk and Schneider's (2002) description of social capital in schools. When teacher A decides to stay after school to help teacher B with the new math curriculum, teacher B benefits from the social capital generated by teacher A's organizational citizenship behavior.

While the procedural justice literature has established that for a decision making process to contribute to social capital it must be perceived as fair (Colquit et al., 2013), it is less clear about *how* decisions should be made fairly. The quality of the decision is left

to the idiosyncrasies of the decision-maker, the context, and those affected by the decision. Here we take a more systemic approach, arguing that decisions will be perceived as fair, and the authority of the decision-maker legitimate, when they are part of a system of governance that balances the claims of different stakeholders.

Figure 1 illustrates our logic model. Consider social capital flows in the middle of the figure. Moving to the right, social capital contributes to teachers' ability to employ effective instructional practices (Penuel et al., 2009), innovate (Coburn & Russell, 2008; Frank, Zhao & Borman, 2004), and coordinate (Bryk & Schneider, 2002). In turn, these teacher behaviors contribute to student outcomes (e.g., Nye, Konstantopolous & Hedges, 2004).

Our study then considers the causes of social capital flows, moving to the left of social capital flows in Figure 1. In particular, we argue that teachers benefit from social capital resource flows when others provide resources as manifest through organizational citizenship behaviors, including those behaviors reflecting teacher commitment. Continuing to the left in Figure 1, educators are inclined to engage in organizational citizenship behaviors when they perceive that decisions in their organization are made fairly and by authorities perceived to be legitimate. Finally, our study focuses on how a system of governance that balances the interests of stakeholders can contribute to decisions that are perceived to be fair, and decision makers perceived as legitimate. These are the arrows and boxes highlighted in bold. We do not mean to say that each arrow indicates the only cause of a construct. But Figure 1 highlights a logical chain between systemic governance and student outcomes.

Insert Figure 1 here

The Balancing Voices Framework (BVF) of School Governance

Preamble

Our system of governance is about the processes for making decisions in schools. It is not about the substance of those decisions themselves. Clearly the curriculum (e.g., Schmidt et al., 2015), pedagogy (e.g., Ball & Bass, 2000), and how schools and learning are structured (e.g., Blumenfeld, et al., 1991; Elmore, Peterson & McCarthy, 1996; Krajcik & Blumenfeld, 2006) can have direct effects on how teachers teach individually and collectively to improve student outcomes. Here we recognize the vast literature that can inform schools in each of these areas as well as the prerogative necessarily exercised by local administrators and leaders (Spillane & Thompson, 1997; Spillane and Zeuli, 1999). Our focus is then on the *process* by which decisions are made about curriculum, pedagogy, and the structuring of the school and learning.

Goals

The BVF supports fair and legitimate decision-making by formally specifying how the voices of teachers, administrators, and community members are balanced in decisions concerning the adoption of reforms and the evaluation of personnel. The BVF consists of 5 goals, presented with a principle for action and examples. See Technical Appendix A for an overview to the detailed descriptions below.

Goal 1. Enhance educators' investment in, and commitment to, reforms to ensure success and sustainability. The principle of action is that even promising, high quality reforms imposed without educator commitment and investment tend to meet resistance, apathy or other responses that reduce fidelity and effectiveness (e.g., Bryk et al., 2015a, 2015b; Fullan, 2010, 2011). At the individual level, lone wolves, who choose their own practices independent of the school's coordinated goals are typically less effective than teachers who coordinate with others in their schools (Harris, Ingle, & Rutledge, 2014). To enhance educators' collective investment in and commitment to reforms we give administrators and teachers a formal voice in the adoption of reforms. For example, similar to the procedures used to implement Success for All, the collective voice of the teachers could be expressed through a formal vote (Borman & Hewes, 2002). But it could also be expressed through a systematic survey or a representative committee.

Goal 2. Allow time for educators to balance fidelity of implementation of reform with local adaptation through collaboration. The principle of action is that it takes time for educators to comprehend and adapt reforms to their settings, and then to re-establish coordination once they have done so (Bryk et al., 2015a, 2015b; Coburn, 2003; Cook et al., 2000; Frank et al., 2011; Fullan 2010, 2011). Extreme and immediate pressure to improve test scores may discourage educators from adopting particular innovations (Frank, Zhao & Borman, 2004). One way to give school professionals adequate time to comprehend, adapt and implement reforms would be to specify that a reform or innovation would not be evaluated using standardized tests for at least three years.

We acknowledge that schools should be held accountable, and that standardized tests may be part of that accountability, even on a yearly basis (Harris & Herrington, 2006; Weixler, Harris and Barrett, 2017). But specific *reforms* should not be expected to improve standardized test scores in the first two years of implementation. Educators should express this expectation to the community at board meetings, through written communications, and, most importantly, as teachers engage parents in routine interactions. The same principle implies that a school should not introduce new or additional reforms that might compete with existing reforms until existing reforms have had the opportunity to mature and can reasonably be summatively evaluated (Borman et al., 2003; Wilkes & Bligh, 1999). Throughout, *formative* evaluations of a reform, including its effects on student and teacher motivation, engagement and collaboration, can and should be conducted in the first years of implementation (e.g., Bloom, Hastings & Madals, 1971).

Goal 3: Coordinate community voice in school affairs. The principle of action is that competing or conflicting pressures may result in a fragmented school, for example as teachers compete for student enrollments in their preferred courses (Bidwell, Frank & Quiroz, 1997; Powell, Farrar & Cohen, 1985). Given that community members outside the school are often not privy to the internal decisions of a school, we conceive their involvement as occurring primarily through a formal expression used in the evaluation of school administrators. As an example, community voice could be expressed through a local school council (e.g., Bryk et al.; 1998, 2010) or through a formal survey regarding building administrators. These formal expressions of community voice can then be

incorporated into a portfolio of materials used in the evaluation of building administrators.

Goal 4: Enhance formal voice for teachers concerning school administrators.

Teachers' power in school governance has eroded with the decline of unions and with lack of recognition of teaching as a profession (Ingersoll & Perda, 2008; Moe, 2010). The principle of action is that without input, teachers will not perceive top-down, administrative decisions as fair or legitimate, and in turn schools may experience declining teacher investment in and commitment. The Balancing Voices framework directly addresses the power dynamic within a school by providing the opportunity for teachers to formally express their voice concerning their administrators. Formal voice could be directly exercised through a union representative, a formal confidence/no confidence vote on the principal, an organized petition expressing support (or not) administration, or a systematic survey addressing the administrators' job performance. Like the community formal voice, the formal voice of teachers could be incorporated into the portfolio of materials used to formally evaluate the administration.²

Goal 5: Encourage ongoing teacher improvement and quality with due process.

For teachers performing below an established threshold (e.g., those deemed "ineffective"), the BVF engages third parties such as an instructional coach or union representative to improve teacher quality and ensure due process should discipline or

² An unusually high rate of teacher turnover could also indicate concerns about administrators, although teachers may have limited capacity to leave their schools in the current climate of education policy.

discharge ensue. The principle of action is that current procedures for addressing underperforming teachers do little to engage the school as a community. As a result, they often do not result in improvement, while at the same time it is difficult to remove an underperforming teacher. The third party would occupy a mediating role between administrators and teachers, ensuring that administrators adhere to specified procedures for evaluating teachers, supporting their improvement, and in some cases participating in the final steps of the dismissal process. For example, a coach or union representative can serve as the third party to the development and implementation of a teacher improvement plan. The presence of a third party should signal the necessity for improvement and give legitimacy to the process, thereby eliciting commitment from teachers and increasing the chances of improvement. If a teacher fails to improve, the third party can verify the extent to which the required steps towards improvement were fulfilled, which would mitigate against a later claim that due process was denied.

Integrated Goals

The integrated BVF is shown in Figure 2. Goals 1 and 2 together help educators avoid potentially polarizing whipsaw responses to competing demands for multiple reforms generated at the federal, state, or district levels. Goal 3 then provides for direct community involvement in school decision-making. Goals 4 and 5 complement one another. In particular, if an administrator moves to discipline or dismiss a teacher perceived by other teachers to be ineffective, the other teachers may choose *not* to formally express a concern about the administrator. If no collective action is pursued, teachers' inaction implies support for the administrator, without forcing the teachers to

explicitly express their concerns about a colleague. If an administrator instead acts against a teacher whose peers perceive to be effective, teachers have a formal mechanism to express disapproval. In either circumstance, the BVF draws on the collective knowledge of the educators about one another (Grissom & Loeb, 2017; Harris, Ingle, & Rutledge, 2014; Jacob & Lefgren, 2008), and of the general value of peer-to-peer evaluation for professional adults (see the review in Gill, Lerner & Moesky, 2015).

These goals complement the existing powers of school boards to make personnel decisions and of communities to elect school board members. Thus, together with existing legal powers to elect boards that can dismiss educators, the BVF completes the set of checks and balances among the stakeholders. See technical appendix B for an analysis of the legality of the BVF in Michigan.

Insert Figure 2 here

Mitigating External Forces

The checks and balances in the BVF allow schools to mitigate the effect of external forces that might otherwise diminish social capital. By requiring that schools implement reforms collectively and with commitment, Goals 1 and 2 provide a school with a buffer against special interests or focused forces that can factionalize a school through targeted but competing demands (Frank, Penuel et al., 2013; Honig & Hatch, 2004). By channeling parental input, Goal 3 allows educators to work collectively as they respond to a coherent community voice. Goals 4 and 5 provide for accountability

through the collective and local knowledge of the teachers and administrators, in contrast to an individualistic or mechanical value-added score (e.g., teacher level value added, see Goldhaber, 2015). This creates an incentive for teachers to contribute to the school as a whole instead of focusing exclusively on the measured achievement levels of only their students (perhaps even by lobbying for particular students, see Kim, Frank & Spillane, 2018; Lankford, Loeb & Wyckoff, 2002). As a set the goals facilitate the commitment of teachers and administrators to one another, making the school an attractive place to work and thus retaining educators in light of external labor market forces (Grissom, 2011).

Beyond Formal Processes

Although the BVF explicitly addresses the formal processes for making decisions about reforms and personnel, the implications go beyond those formal processes. For example, we anticipate the BVF would generate extensive deliberation, thereby catalyzing honest and direct discussions among stakeholders, and increasing the likelihood that all will perceive the ultimate process as fair and legitimate (e.g., Briggs, Russell & Wanless, 2018). We also anticipate that by engaging all stakeholders formally and systematically in decisions of the school, the BVF would make decisions more transparent than conventional practices, again contributing to the perceived fairness and legitimacy of the process. Finally, the formal processes of the BVF will encourage uniform access to input in decision-making, thereby discouraging selective access of the “squeaky wheels.”

Note the emphasis on processes in the preceding paragraph. In fact, in an ideal setting the formal rules or policies associated with the BVF might never be invoked. For example, if all stakeholders know that educators will have formal voice in a particular decision then educators will be engaged throughout the process of that decision. That engagement itself may shape the decision, regardless of a final formal expression of the educators. Similarly, the formal capacity of the community to express concerns about an administrator could generate forthright and constructive communication between parents and administrators regardless of the final formal evaluative expressions of the parents.

Although we have good reason to believe that the BVF would have beneficial effects in full implementation and practice, such implementation would take years to evaluate (Goal 1) and likely be expensive (Slavin, 2008).³ Therefore, to fairly and economically test the key elements of the BVF, we turn to a randomized control trial in a simulated environment.

Methods

The study consists of role-play simulations in groups of 4-5 in which participants were randomly assigned to use a protocol based on the BVF or business as usual. After the simulations participants completed a short survey.

³ We are currently engaged with 3 districts who are considering implementing the BVF.

Sample

Our sample consists of 122 participants from Michigan who were graduate students in education or working in the educational policy arena. We recruited our participants in six settings from courses in educational policy or leadership, and in an educational policy fellows forum (<http://epfp.educ.msu.edu/>). About 77% of the participants were former or current school teachers or administrators indicating formal engagement with the school and policy contexts.

Procedures

In each setting at least three of the four co-authors modeled a role play with the following scenario:

There's new legislation that allows the district to determine the school schedule. A principal wants to adopt a year-round schedule (drawing on some evidence supporting this especially for disadvantaged children). Setting: suburban high school.

The demonstration role play was conducted both with a Balancing Voices protocol and a business as usual protocol. This was intended to familiarize the participants with the nature of the roles (e.g., principal; teacher with students with mediocre test scores based on value added; teacher with students with low test scores based on value added; community member; district administrator) and the process of the simulation.

Participants in each setting then were randomly assigned to groups of 4-5 based on a draw of playing cards (e.g., for two groups: black face cards, red face cards). Having participants draw the cards themselves made the randomization transparent.

Originally there were 30 groups across the five settings, but due to a clerical error some groups within a common setting were consolidated resulting in 23 groups used for analyses.⁴

Members of each group then assigned themselves to fill the roles while enacting a given scenario. Allowing the participants to choose their own roles contributed to authenticity, while the initial randomization ensured that there were no expected differences between the groups.

Each group was then assigned one of the following two scenarios to discuss and attempt to come to a resolution:

1: A principal would like to remove a teacher whose has consistently low test scores (value-added). The teacher has one or two close colleagues in the school.

Setting: urban high school.

Assumption: teacher knows principal does not like him/her

⁴ We were not able to differentiate the group memberships between 7 pairs of groups that occurred within two rounds of data collection. In these cases only we analyzed all those we could not differentiate as though they were in common groups of Balancing Voices or “business as usual.” This produced two treatment groups of size 12 and 8 and one control group of size 14. Our analytic approach is conservative as our main results are based on a multilevel model with degrees of freedom defined at the group level. But given relatively small intra-class correlations (less than 10% for each outcome), standard errors are only minimally responsive to the number of groups and therefore to this decision.

2: *The teachers are divided on pedagogy (e.g., 50% seek to implement a new reform and 50% prefer current practices) and the principal is not prone to action, but is well supported by the district.*⁵

Setting: suburban elementary school

Participants in black card groups were instructed to use a “business as usual” protocol (as defined in Michigan) for making decisions. Participants in red card groups were asked to use a specific protocol derived from the BVF. The majority of participants were accustomed to business as usual rules within the local state context. The differences in the dependent variables by protocol are shown in Table 1 below. We do not report significance levels because these values are at the individual level while the treatment was implemented at the group level. Note that the difference favors Balancing Voices by an effect size of about .38 ($.25/.65=.38$) for procedural fairness, about .57 ($.4/.7=.57$) for legitimacy of authority, and about .25 ($.16/.65=.25$) for outcome favorability (effect sizes calculated using the pooled standard deviation based on Table 1).

Insert Table 1 here

⁵ Scenario suggested by Rachel Fish.

Each group was given approximately 20 minutes to discuss the assigned scenario. After discussing the scenario, the participants were asked to complete a short (approximately five minute) survey including background characteristics, a fidelity check, and measures of outcome favorability, procedural fairness and legitimacy of authority figures.

Twenty-four people in one particular setting participated in two rounds, one for each scenario, with groups randomly reassigned between rounds. Unconditional models accounting for repeated measures on these participants indicated that there was not statistically significant (for $\alpha=.05$) variation due to the repeated measures for any of our outcomes.

Measures

Fidelity check. Respondents indicated the protocols their group used by identifying which rule of a pair they had used for particular decisions. Details are provided in Technical Appendix C.

Dependent Variables. Following the procedural fairness literature (e.g., Van der Toorn, Tyler, & Jost, 2011), we adapted our dependent measures to our context (Colquitt, 2001; Colquitt et al., 2013). All were based on Likert scales: 1=strongly disagree, 2=disagree, 3=agree, 4=strongly agree.

Procedural fairness. We used a direct measure of procedural fairness because our measure used the term “fairness” (Colquitt et al., 2013). In particular, our items were: I noticed that some people/groups were treated unfairly (reverse coded); the decision was reached in a fair and equitable manner; it is fair to decide issues like the one today in the manner we did ($\alpha=.70$).

Legitimacy of authority figures. The procedural fairness literature notes a subtle distinction between the decisions and those (typically supervisors) who made them (see Colquitt et al., 2013, page 203).⁶ In order to capture group members’ perception of their decision makers, we measure the perceived legitimacy of authority figures. This was based on two items (correlated at .68): I found the principal in the simulation trustworthy; I believe the principal placed a high priority on the school’s best interest.

Outcome favorability. Our measure of outcome favorability was based on the mean of two items (correlated at .35): I found the outcome personally favorable; I was frustrated by today’s decision (reverse coded).

Covariates. Because participants were randomly assigned to Balancing Voices and business as usual our estimates of the Balancing Voices effect are unbiased regardless of whether we control for covariates. But we explore covariates primarily so

⁶ Colquitt et al. (2013) found no differences in effects between perceptions of procedures and of supervisors (see page 208).

that we may examine whether there are heterogeneous effects of Balancing Voices by role or scenario.

Role. We recognize that the role a participant played might have affected her ratings. Therefore, we created indicators of the role played (e.g., principal, teacher with students with low test scores, teacher with students with mediocre test scores; teacher who supports a reform; community member; district administrator; other).

Scenario. The content of the scenarios may have affected participants' responses to the decisions and processes. Therefore we created an indicator for scenario (taking a value of 1 for the first scenario for the removal of a teacher, 0 for the teachers with divided pedagogy).

Analytic Approach

We first evaluated the fidelity of adherence to the assigned decision-making protocol. We then generated descriptive statistics by protocol and examined boxplots of the outcomes by protocol. We then estimated the effect of Balancing Voices. Since the treatments were assigned at the group level, we used a multilevel model to further test differences (Raudenbush & Bryk, 2002). For example, for participant i in group j for procedural fairness we modeled:

Level 1 (participant i):

$$\text{Procedural fairness}_{ij} = \beta_{0j} + e_{ij},$$

where the e_{ij} are assumed iid normal $(0, \sigma^2)$.

For level 2 (the group j):

$$\beta_{0j} = \gamma_{00} + \gamma_{01} \text{Balancing Voices}_j + u_{0j},$$

where the u_{0j} are assumed iid $N(0, \tau_0)$. The variable $\text{Balancing Voices}_j$ is an indicator of whether the group members were randomly assigned to the Balancing Voices protocol or the conventional “business as usual” protocol. Correspondingly, the term γ_{01} indicates any resulting difference in perception between the Balancing Voices groups and the conventional rules groups. We then added the role of the participant through a set of indicators at level 1 and scenario through an indicator at level 2. We also examined cross-level interactions between Balancing Voices and role.

We estimated our models using SAS prox mixed with $\text{ddfm} = \text{BW}$, the more conservative approach outlined by Singer (1998).⁷ We also verified our inferences using sandwich estimators of standard errors (Huber, 1967; White 1980, 1994). Recognizing that our inferences may be due to bias generated by measurement error, Hawthorne effects, lack of fidelity to treatment assignment, our particular sample, or other mechanisms, we quantified how much bias must have been present to invalidate any inferences we made (Frank et al., 2013).

As an exploratory analysis, we also investigate whether the estimated effect of Balancing Voices is stronger for those who perceived outcomes of the discussion to be

⁷ Inferences identical using $\text{df} = \text{Kenward-Roger}$ or using a sandwich estimator with the empirical option (which actually produces smaller standard errors).

less favorable (the models include controls for scenario, role, and person random effects). An effect in this direction would imply that Balancing Voices can be especially valuable in helping organizations maintain commitment of even among those who do not benefit directly from a particular decision.

Results

Descriptive statistics are presented in Table 2 and box plots of each outcome by decision-making protocol are shown in Figures 3a,3b and 3c. Visual inspection shows that the Balancing Voices protocol generated considerably higher responses than the conventional rules, with the marker for the bottom quartile for Balancing Voices aligning with the median for conventional rules across for procedural fairness and legitimacy of authority.

Insert Table 2 here

Insert Figure 3 here

Estimates of the effect of the Balancing Voices protocol are shown in Table 3. In the first column of Table 3, the unadjusted estimated effect of Balancing voices are consistent with Table 2, with a .25 difference for procedural fairness ($p \leq .05$) and a .5 unit difference for legitimacy of authority figures ($p \leq .01$). The difference on outcome favorability was not statistically significant, suggesting that the effect of Balancing

Voices was not necessarily due to producing more favorable outcomes, but instead due to a more legitimate process. These significance levels and corresponding inferences do not change across the columns when different controls are included in the model. There were not significant interactions between Balancing Voices and role. The estimated Balancing Voices effect was stronger and borderline statistically significant across all outcomes ($p < .05$ only for legitimacy of authority figures) for the second scenario (regarding the teachers with divided pedagogy).

Insert Table 3 here

Robustness of the Inferences

Using the approach of Frank et al. (2013) as applied in Rosenberg, Xu and Frank (2018), 32% of the estimated effect of Balancing Voices of legitimacy of authority figures from the last model in Table 3 with all controls would have to be due to bias to invalidate the inference (using the Konfound-it.com app with estimated effect of .32, standard error of .11, sample size of 123 and number of covariates=13). A similar % would apply to the model with no controls. Note that the controls for role, scenario, and participants included in the final model only accounted for a change of estimated effect of .08, from .40 to .32. To invalidate the inference, other sources of bias would have to reduce the estimate by another .10, conditional on the covariates already in the model. That is, the omitted sources of bias would have to be stronger than the sources of bias accounted for by role, scenario, and participant effects. Interpreting in terms of case replacement (Frank et al., 2013), to invalidate the inference of an effect of Balancing Voices on legitimacy of authority figures, 32% (39) of the cases would have to be replaced with null hypothesis cases in which Balancing Voices had no effect.

Using similar calculations, 21% of the estimated effect of Balancing Voices on procedural fairness would have to be due to bias to invalidate the inference of an effect of Balancing Voices on procedural fairness (about equal to the % change in the estimate due to the controls we used in the model). The median value for % bias to invalidate an inference reported on in Frank et al. (2013) was about 30%. Thus, our inferences are at least moderately robust to concerns about measurement error, Hawthorne effects, sample representativeness and statistical procedures.

Interaction with Outcome Favorability

Table 4 includes separate models estimated for procedural justice and legitimacy of authority using outcome favorability as a predictor as well as interacting outcome favorability with balancing voices. Not surprisingly, the main effect of outcome favorability on legitimacy of authority and procedural fairness is strong and positive. When participants felt the outcome favored them they were more accepting of the process. Interestingly, however, the interaction between Balancing Voices and outcome favorability is negative and approaches statistical significance. That is, Balancing Voices may have had a modestly stronger effect for those who perceived that the outcomes did not favor them. While this is exploratory, if the result pertained it would suggest that Balancing Voices may help organizations maintain commitment even among those who do not directly benefit from particular decisions.

Insert Table 4 about here

Discussion

While schools with high levels of social capital and educator commitment to the organization may perform better, we know little about the actions that schools and districts can take to cultivate social capital and commitment. Here we examine how social capital can be cultivated by formalizing the processes through which schools make fundamental decisions regarding reforms and how they evaluate teachers and principals using the BVF. In this sense, we link school governance to student outcomes through social capital and educator commitment.

To test our theory about the relationship between Balancing Voices and procedural fairness we developed a specific operationalization of the BVF. We then conducted a randomized experiment with role-play simulations. We found that those assigned to rules derived from the BVF perceived greater legitimacy for their group's decision and the administrator making it than did those assigned to "business as usual" processes (as in Michigan).

Interestingly, the participants did not perceive *outcomes* to be significantly more favorable under the BVF than under business as usual. Thus, we cannot, and do not, claim that the BVF will lead to better decisions in any particular context. Instead, we claim the BVF leads to a better decision-making *process*.

While a focus on process may seem indirect, effective processes address the challenge of building organizational culture. Commitment to the school is through the process, which creates the expectation for future allocations. Thus, a single process can enhance the commitment of *all* educators to the school, even those who did not benefit from an immediate resource allocation (e.g., the teacher deemed underperforming at a

given point in time). In fact, the test of an effective process is its ability to sustain commitment even from those who did not immediately benefit from the decision. Thus it is the process that ultimately gives legitimacy to the school and those making the decisions in the school.

Limitations

Single state sample. Since our sample is from a single state, the participants in our experiment may not represent those in other states. Moreover, the “business as usual” rules we used were based on rules in Michigan. Other states clearly have different bases for removing teachers and principals, as well as different levels of decentralization for making decisions about the implementation of reforms. Even so, Michigan is not extreme in its approach to the evaluation of reforms and educators. Like Michigan, many states have a theoretical right to work or broad management rights in principle, but in practice the removal of a teacher is costly and time intensive (e.g., Bridges, 1990; Stoelinga, 2017). Furthermore, our results are fairly robust to moderate deviations in the sample that might include somewhat weaker effects in different settings.

Regarding the decentralization of decision-making, Michigan well represents the current state of uncertainty. Generally, federal and state governments attempted to tighten organizational coupling in the early 2000’s by holding schools accountable through the use of standardized tests, and later through the use of federally supported curricular standards (Boyd & Crowson, 2002; Cohen & Moffitt, 2009). But since 2015, federal and national efforts to tighten coupling have been relaxed as the federal government has given states more flexibility. Recently, federal ESSA legislation reduced

emphasis on the use of standardized tests for evaluation, particularly for teachers (Klein, 2016). Across states, legislation has modified or rejected elements of the curricular standards such as the Common Core State Standards (e.g., Layton, 2014). In tandem, recent changes may have returned some autonomy to states, districts, and schools (Klein, 2016).

Only one interpretation of the BVF tested. The rules we used in our simulations are but one operationalization of our framework. On the one hand, other interpretations of the framework may not generate exactly the same results we observed here. On the other hand, the value of the framework is that it can be easily adapted to various settings. For example, some schools might replace formal votes with systematic surveys already in place. These would be consistent with the expression of voice within the framework, even though they were not explicitly tested here. In any case, while districts may choose their own interpretations of the framework, any interpretation must be consistent with the holistic of balancing voices the framework instantiates. Toward that end, the study here yields encouraging evidence from the first test of such a holistic approach to governance.

Conclusion

Schools' and districts' may have responded slowly to their newfound autonomy because doing so could upset previous distributions of power. Therefore, by formalizing the balance of power, the BVF can support schools and districts seeking more autonomy afforded them under the ESSA.

Having established the potential of one instantiation of the BVF in a laboratory setting, we hope that schools will now be emboldened to generally reflect on their governance structures, and more specifically, to consider a governance approach that balances the voices of their key stakeholders. Given the nature of policy windows in general (Kingdon & Thurber, 1984) we expect that the current opportunity for schools and districts to assert their autonomy will not be sustained as federal and state policymakers seek to reassert control that serves *their* policy and political goals. The time for schools and districts to seize decision-making control is likely now.

Technical Appendix A
Overview of Goals, Principles of Action and Examples of Balancing Voices

Goal	Principle of Action	Examples
<i>Enhance educators' investment in, and commitment to, reforms to ensure success and sustainability</i>	Reforms imposed without educator commitment and investment tend to meet resistance, apathy or other responses that reduce fidelity and effectiveness	Formal vote of teachers regarding reform adoption; Systematic survey of teachers and administrators
<i>Allow time for educators to balance fidelity of implementation of reform with local adaptation through collaboration</i>	It takes time for educators to comprehend and adapt reforms to their settings, and then to re-establish coordination once they have done so.	Minimum three-year period after implementation begins, during which a reform or innovation is not evaluated using standardized tests
<i>Coordinate community voice in school affairs</i>	When community input is not formally organized, it can be limited or manifest in competing pressures within a school.	Local school council; Formal survey of community members
<i>Enhance formal voice for teachers concerning school administrators</i>	Without input regarding administrators, teachers will not perceive top-down, administrative decisions as fair or legitimate, and in turn schools may experience declining teacher investment and commitment to the school.	Union representative; Formal vote of confidence/no confidence; Formal petition; Formal survey
<i>Encourage ongoing teacher improvement and quality with due process</i>	The third party occupies a mediating role between administrators and teachers, ensuring that administrators adhere to specified procedures for evaluating teachers, supporting their improvement, and in some cases ultimately dismissing them.	A coach or union representative can serve as the third party to the development and implementation of a teacher improvement plan

Technical Appendix B

Legality of the Balancing Voices Goals in Michigan.

The BVF may challenge existing business as usual norms in districts or state education legislation. We use Michigan as a case study to examine the Framework’s alignment within existing school code. Table B1 below presents Michigan legislation and BVF alignment.

Table B1: Michigan legislation and the BVF

Legislation	Oversight	Balancing Voices
Public Employment Relations Act 336 of 1947	Schools may determine the procedures for decisions concerning experimental or pilot programs	Goal 1: The majority of instructional staff must agree to a new reform prior to its enactment
		Goal 2: Prior to evaluation, schools must be given three years to enact a new educational reform
	The school may determine the <i>procedures</i> used for implementing evaluation	Goal 3 and 4: Community and teacher engagement may contribute a vote of confidence to the evaluation the school principal
PA 173 Section 1249b (2.1.d.iv)	Annual administrator evaluations must include feedback from the school community which includes teachers, students, and parents.	Goals 3 & 4 explicitly provide an avenue for community and teacher voice regarding school leadership.
PA173 Sec.1249 (2)	For teachers rated as minimally effective or ineffective), the school administrator or designee shall develop, in consultation with the teacher, <i>an individualized development plan that includes these goals and training and is designed to assist the teacher to improve his or her effectiveness.</i>	Goal 5 designs a teacher evaluation and improvement plan with multiple checkpoints for improvement in alignment to state law.

Based on conversations with a hired lawyer who works in teacher labor and collective bargaining in Michigan, we have assessed the legality of the Balancing Voices goals.

Regarding goal 1, section 15 part 3 (c) of the Public Employment Relations Act 336 of 1947 states that the school may determine the *procedures* for decisions concerning use and staffing of experimental or pilot programs. Thus, the school may choose the procedures for assigning teachers to reforms. Goal 2 pertains more to the criteria a district uses to evaluate itself internally, and so does not directly engage state law.

Regarding goals 3 and 4, Act 336 of 1947 states: “The school may determine the *procedures* used for implementing evaluation.” Specific to goal 3, the PA173 Sec.1249b (2.1.d.iv) concerning building level administrator evaluation should include “Student, parent, and teacher feedback, as available, and other information considered pertinent by the superintendent or other school administrator conducting the performance evaluation or the board or board of directors.” Thus the community (as in goal 3) and teachers (as in goal 4) have a rightful place in the evaluation system.

Specific to goal 5, the PA173 Sec.1249 (2) states: For a teacher described in subdivision (d) (added: which includes teachers rated as minimally effective or ineffective), the school administrator or designee shall develop, in consultation with the teacher, *an individualized development plan that includes these goals and training and is designed to assist the teacher to improve his or her effectiveness*. Thus goal 5 complies with Michigan State law.

Technical Appendix C
Fidelity Check

The pairs are presented below in the order they appeared on the survey. The bold element in each pair represented the BVF:

Reform Adoption:

The state or district determined the adoption of reform.
The principal and the teachers voted on reform.

Evaluation of reform:

A reform is not evaluated using standardized tests until the reform has been implemented for three years.

The success of a new curriculum or reform is evaluated in terms of changes in test scores after it is implemented.

Evaluation and dismissal of a principal:

A principal’s evaluation includes a vote of the majority of the local school council and an option for teachers to express concern by signing a petition or submitting for leave.

A district superintendent can fire or remove a principal at the end of a school year because the principal is ineffective or negligent.

Evaluation and dismissal of a teacher:

Due process (3 years of documentation) is required for a principal to remove a teacher.

The principal has an expedited path to dismissing teachers. The primary union representative will participate throughout the expedited evaluation process, providing a signature of approval indicating the process is being fairly adhered to and implemented.

As shown in Table C1, overall the fidelity was high with small percentages (4%-28%) of the respondents deviating from the assigned rules for governance.

Table C1: Fidelity to Assigned Protocols

Item	Conventional (n=53)	Balancing Voices (n=52)
<i>Reform Adoption</i>	84%	96%
<i>Evaluation of reform</i>	89%	72%
<i>Evaluation and dismissal of a principal</i>	86%	90%
<i>Evaluation and dismissal of a teacher</i>	79%	96%

19-22 cases missing depending on the item

References

- Ball, D. L., & Bass, H. (2000). Interweaving content and pedagogy in teaching and learning to teach: Knowing and using mathematics. In J. Boaler (Ed.), *Multiple perspectives on mathematics teaching and learning* (pp. 83-104). Westport, CT: Ablex Publishing.
- Bidwell, C. E., Frank, K. A., & Quiroz, P. (1997). Teacher Types, Workplace Controls, and the Organization of Schools. *Sociology of Education*, 70(4), 285-307.
- Bidwell, C. E., & Kasarda, J. D. (1980). Conceptualizing and measuring the effects of school and schooling. *American Journal of Education*, 88(4), 401-430.
- Bloom, B. S., Hastings, J. T., & Madals, G. F. (1971). *Handbook on formative and summative evaluation of student learning*. New York, NY: McGraw-Hill.
- Blumenfeld, P. C., Soloway, E., Marx, R. W., Krajcik, J. S., Guzdial, M., & Palincsar, A. (1991). Motivating project-based learning: Sustaining the doing, supporting the learning. *Educational Psychologist*, 26(3-4), 369-398.
- Borman, G., & Hewes, G. (2002). The long-term effects and cost-effectiveness of Success for All. *Educational Evaluation and Policy Analysis*, 24(4), 243-266.
- Bridges, E. M. (1990). *Managing the Incompetent Teacher*. Eugene, OR: University of Oregon Press.
- Briggs, J. O., Russell, J. L., & Wanless, S. B. (2018). Kindergarten Teacher Buy-In for Standards-Based Reforms: A Dynamic Interplay Between Professional Identity and Perceptions of Control. *Early Education and Development*, 29(1), 125-142.
- Bryk, A. S., & Schneider, B. (2002). *Trust in schools: A core resource for improvement*. New York, NY: Russell Sage Foundation.
- Bryk, A. S., Gomez, L. M., Grunow, A., & LeMahieu, P. G. (2015a). *Learning to Improve: How America's Schools Can Get Better at Getting Better*. Cambridge, MA: Harvard University Press.
- Bryk, A. S., Gomez, L. M., Grunow, A., & LeMahieu, P. G. (2015b). Breaking the Cycle of Failed School Reforms: Using Networked Improvement Communities to Learn Fast and Implement Well. *Harvard Education Letters*, 31(1).
- Bryk, A. S., Pennie, B., Sebring, D. D., Kerbow, S., Rollow, S., & Easton, J. Q. (1998). *Charting Chicago School Reform: Democratic Localism as a Lever for Change*. Boulder, CO: Westview Press.
- Bryk, A. S., Sebring, P. B., Allensworth, E., Easton, J. Q., & Luppescu, S. (2010). *Organizing Schools for Improvement: Lessons from Chicago*. Chicago, IL: University of Chicago Press.
- Coburn, C. E. (2003). Rethinking scale: Moving beyond numbers to deep and lasting change. *Educational Researcher*, 32(6), 3-12.
- Coburn, C. E., & Russell, J. L. (2008). District policy and teachers' social networks. *Educational evaluation and policy analysis*, 30(3), 203-235.
- Coburn, C. E., Russell, J. L., Kaufman, J., & Stein, M. K. (2012). Supporting sustainability: Teachers' advice networks and ambitious instructional reform. *American Journal of Education*, 119(1), 137-182.
- Colquitt, J. A. (2001). On the dimensionality of organizational justice: A construct validation of a measure. *Journal of Applied Psychology*, 86(3), 386-400.

- Colquitt, J. A., Conlon, D. E., Wesson, M. J., Porter, C. O., & Ng, K.Y. (2001). Justice at the millennium: A meta-analytic review of 25 years of organizational justice research. *Journal of Applied Psychology, 86*(3), 424-445.
- Cook, T. D., Murphy, R. F., & Hunt D. H. (2000). Comer's School Development Program in Chicago: A Theory-Based Evaluation. *American Educational Research Journal, 37*(2), 535-597.
- Cohen, D. K., & Moffitt, S. L. (2009). *The ordeal of equality: Did federal regulation fix the schools?* Cambridge: Harvard University Press.
- Elmore, R. F., Peterson, P. L., & McCarthy, S. J. (1996). *Restructuring in the classroom: Teaching, learning, and school organization*. San Francisco, CA: Jossey-Bass Inc.
- Folger, R., & Konovsky, M. A. (1989). Effects of procedural and distributive justice on reactions to pay raise decisions. *Academy of Management Journal, 32*(1), 115-130.
- Frank, K. A., Kim, C., & Belman, D. (2010). Utility Theory, Social Networks, and Teacher Decision Making. In A. J. Daly (Ed.), *Social Network Theory and Educational Change* (pp. 223-42). Cambridge, MA: Harvard University Press.
- Frank, K. A., Zhao, Y., & Borman, K. (2004). Social Capital and the Diffusion of Innovations within Organizations: Application to the Implementation of Computer Technology in Schools. *Sociology of Education, 77*(2), 148-171.
- Frank, K. A., Zhao, Y., Penuel, W. R., Ellefson, N. C., & Porter, S. (2011). Focus, Fiddle and Friends: Sources of Knowledge to Perform the Complex Task of Teaching. *Sociology of Education, 84*(2), 137-156.
- Frank, K. A., Maroulis, S., Duong, M., & Kelcey, B. (2013). What Would It Take to Change an Inference? Using Rubin's Causal Model to Interpret the Robustness of Causal Inferences. *Education, Evaluation and Policy Analysis, 35*(4), 437-460.
- Fullan, M. (2010). *All systems go: The change imperative for whole system reform*. Thousand Oaks, CA: Corwin Press.
- Fullan, M. (2011). *Change leader: Learning to do what matters most*. San Francisco: Jossey-Bass, Inc.
- Gill, B. P., Lerner, J. S., & Meosky, P. (2015). *Reimagining Accountability in K-12 Education: A Behavioral Science Perspective*. Paper presented at the annual meeting of the Association of Public Policy and Administration, Miami, FL.
- Goldhaber, D. (2015). Exploring the potential of value-added performance measures to affect the quality of the teacher workforce. *Educational Researcher, 44*(2), 87-95.
- Grissom, J. A. (2011). Can good principals keep teachers in disadvantaged schools? Linking principal effectiveness to teacher satisfaction and turnover in hard-to-staff environments. *Teachers College Record, 113*(11), 2552-2585.
- Grissom, J. A., & Loeb, S. (2017). Assessing principals' assessments: Subjective evaluations of teacher effectiveness in low-and high-stakes environments. *Education Finance and Policy, 12*(3), 369-395.
- Grissom, J. A., Nicholson-Crotty, S., & Harrington, J. R. (2014). Estimating the effects of No Child Left Behind on teachers' work environments and job attitudes. *Educational Evaluation and Policy Analysis, 36*(4), 417-436.
- Harris, D. N., Ingle, W. K., & Rutledge, S. A. (2014). How teacher evaluation methods

- matter for accountability: A comparative analysis of teacher effectiveness ratings by principals and teacher value-added measures. *American Educational Research Journal*, 51(1), 73-112.
- Harris, D. N., & Herrington, C. D. (2006). Accountability, standards, and the growing achievement gap: Lessons from the past half-century. *American journal of education*, 112(2), 209-238.
- Holme, J. J., Jabbar, H., Germain, E., & Dinning, J. (2017). Rethinking Teacher Turnover: Longitudinal Measures of Instability in Schools. *Educational Researcher*, 47(1), 62-75.
- Honig, M. I., & Hatch, T. C. (2004). Crafting coherence: How schools strategically manage multiple, external demands. *Educational Researcher*, 33(8), 16-30.
- Hornig, E. L., Klasik, D., & Loeb, S. (2010). Principal's time use and school effectiveness. *American Journal of Education*, 116(4), 491-523.
- Huber, P. J. (1967). The Behavior of Maximum Likelihood Estimates under Nonstandard Conditions. *Proceedings of the Fifth Berkeley Symposium on Mathematical Statistics and Probability*, Vol. I, pp. 221-33.
- Ingersoll, R. M., & Perda, D. (2008). The status of teaching as a profession. In J. H. Ballentine, & J. Z. Spade (Eds.), *Schools and society: A sociological approach to education* (pp.107-118). Belmont, CA: Wadsworth/Thomson Learning.
- Ingersoll, R. M., & Strong, M. (2011). The impact of induction and mentoring programs for beginning teachers: A critical review of the research. *Review of educational Research*, 81(2), 201-233.
- Jacob, B. A., & Lefgren, L. (2008). Can principals identify effective teachers? Evidence on subjective performance evaluation in education. *Journal of Labor Economics*, 26(1), 101-136.
- Kim, C. M., Frank, K. A., & Spillane, J. P. (2018). Relationships among Teachers' Formal and Informal Positions and Their Incoming Student Composition. *Teachers College Record*, 120(3).
- Klein, A. (2016). Under ESSA, states, districts to share more power. *Education Week*, 35(15), 10-12.
- Krajcik, J. S., & Blumenfeld, P. C. (2006). Project-based learning. In R. K. Sawyer (Ed.), *Cambridge Handbook of the Learning Sciences* (pp. 317-34). New York, NY: Cambridge University Press.
- Lankford, H., Loeb, S., & Wyckoff, J. (2002). Teacher sorting and the plight of urban schools: A descriptive analysis. *Educational Evaluation and Policy Analysis*, 24(1), 37-62.
- Layton, L. (2014). Some states rebrand controversial Common Core education standards. *Washington Post* (Jan 30).
- Liang, G., & Akiba, M. (2015). Characteristics of teacher incentive pay programs: a statewide district survey. *Journal of Educational Administration*, 53(6), 702-717.
- Moe, T. M. (2011). *Special interest: Teachers unions and America's public schools*. Washington, DC: Brookings Institution Press.
- Nye, B., Konstantopoulos, S., & Hedges, L. V. (2004). How large are teacher effects? *Educational Evaluation and Policy Analysis*, 26(3), 237-257.
- Organ, D. W. (1988). *Organizational citizenship behavior: The good soldier syndrome*. Lexington, MA: Lexington Books.

- Penuel, W., Riel, M., Krause, A., & Frank, K. A. (2009). Analyzing Teachers' Professional Interactions in a School as Social Capital: A Social Network Approach. *Teachers College Record*, 111(1), 124-163.
- Powell, A. G., Farrar, E., & Cohen, D. K. (1985). *The shopping mall high school: Winners and Losers in the Educational Marketplace*. Boston, MA: Houghton Mifflin.
- Raudenbush, S. W., & Bryk, A. S. (2002). *Hierarchical linear models: Applications and data analysis methods*. San Francisco: Sage.
- Ronfeldt, M., Farmer, S. O., McQueen, K., & Grissom, J. A. (2015). Teacher collaboration in instructional teams and student achievement. *American Educational Research Journal*, 52(3), 475-514.
- Rosenberg, J. M., Xu, R., & Frank, K. A. (2018). Konfound-It!: Quantify the Robustness of Causal Inferences. <http://konfound-it.com>.
- Schmidt, W. H., Burroughs, N. A., Zoido, P., & Houang, R. T. (2015). The role of schooling in perpetuating educational inequality: An international perspective. *Educational Researcher*, 44(7), 371-386.
- Singer, J. D. (1998). Using SAS PROC MIXED to fit multilevel models, hierarchical models, and individual growth models. *Journal of Educational and Behavioral Statistics*, 23(4), 323-355.
- Slavin, R. E. (2008). Perspectives on evidence-based research in education-what works? Issues in synthesizing educational program evaluations. *Educational Researcher*, 37(1), 5-14.
- Smith, C. A., Organ, D. W., & Near, J. P. (1983). Organizational citizenship behavior: Its nature and antecedents. *Journal of Applied Psychology*, 68(4), 653-663.
- Spillane, J. P., & Lee, L. C. (2014). Novice school principals' sense of ultimate responsibility: Problems of practice in transitioning to the principal's office. *Educational Administration Quarterly*, 50(3), 431-465.
- Spillane, J. P., Camburn, E. M., & Pareja, A. S. (2007). Taking a Distributed Perspective to the School Principal's Workday. *Leadership and Policy in Schools* 6(1), 103-25.
- Spillane, J. P., Kim, C. M., & Frank, K. A. (2012). Instructional Advice and Information Providing and Receiving Behavior in Elementary Schools: Exploring Tie Formation as a Building Block in Social Capital Development. *American Educational Research Journal*, 49(6), 1112-1145.
- Spillane, J. P., & Thompson, C. L. (1997). Reconstructing conceptions of local capacity: The local education agency's capacity for ambitious instructional reform. *Educational Evaluation and Policy Analysis*, 19(2), 185-203.
- Spillane, J. P., & Zeuli, J. S. (1999). Reform and teaching: Exploring patterns of practice in the context of national and state mathematics reforms. *Educational Evaluation and Policy Analysis*, 21(1), 1-27.
- Stoelinga, S. R. (2017). Pressuring Teachers to Leave. *Education Week*.
- Sun, M., Loeb, S., & Grissom, J. A. (2017). Building teacher teams: Evidence of positive spillovers from more effective colleagues. *Educational Evaluation and Policy Analysis*, 39(1), 104-125.
- Van der Toorn, J., Tyler, T. R., & Jost, J. T. (2011). More than fair: Outcome dependence, system justification, and the perceived legitimacy of authority figures. *Journal of Experimental Social Psychology*, 47(1), 127-138.

- Weixler, L. B., Harris, D. N., & Barrett, N. (2017). Teachers' Perspectives on the Learning and Work Environments Under the New Orleans School Reforms. *Educational Researcher*, 0013189X18787806.
- White, H. (1980). A Heteroskedasticity-Consistent Covariance Matrix Estimator and a Direct Test for Heteroskedasticity. *Econometrica*, 48(4), 817–838.
- White, H. (1994). *Estimation, Inference, and Specification Analysis*. New York, NY: Cambridge University Press.
- Wilhelm, A. G., Chen, I. C., Smith, T. M., & Frank, K. A. (2016). Selecting Expertise in Context Middle School Mathematics Teachers' Selection of New Sources of Instructional Advice. *American Educational Research Journal*, 53(3), 456-491.
- Wilkes, M., & Bligh, J. (1999). Evaluating educational interventions. *BMJ*, 318, 1269-1272.

Figure 1. Logic Model from Organizational Governance to Social Capital and Student Outcomes

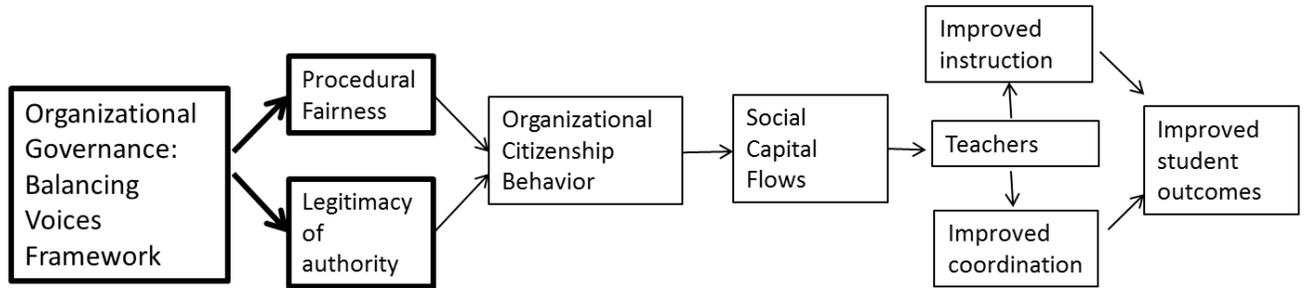
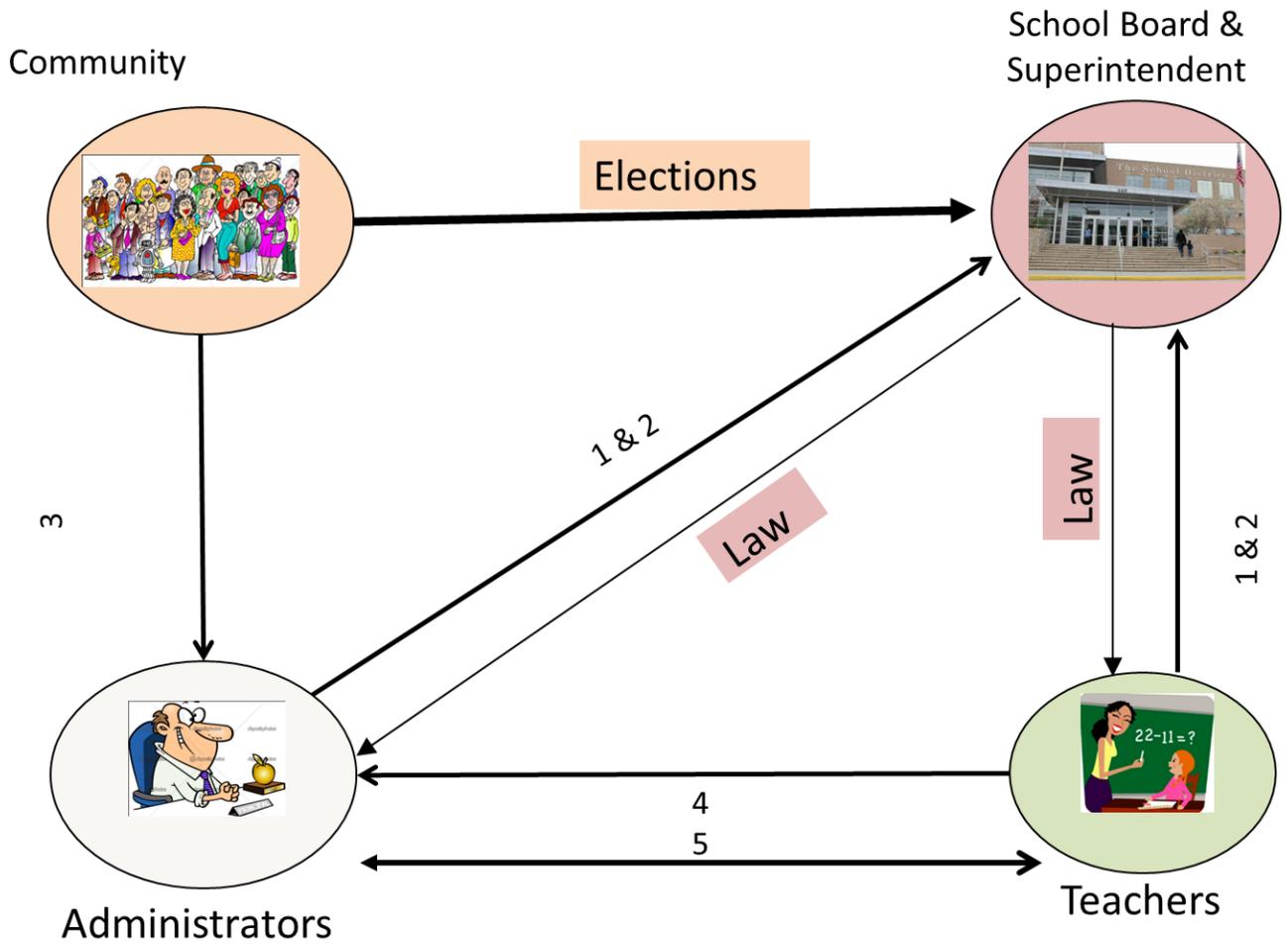


Figure 2: Checks and Balances in the Balancing Voices Framework



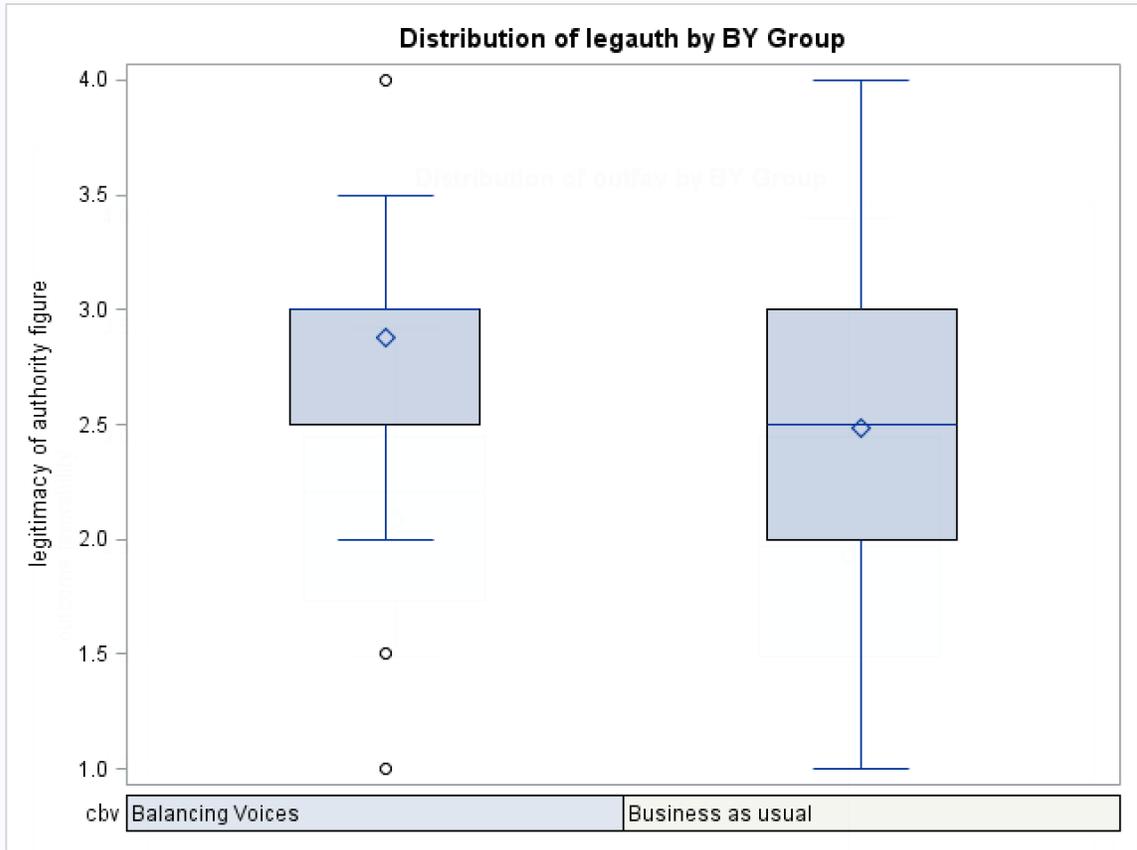


Figure 3a. Boxplots of Legitimacy of Authority Figures by Protocol

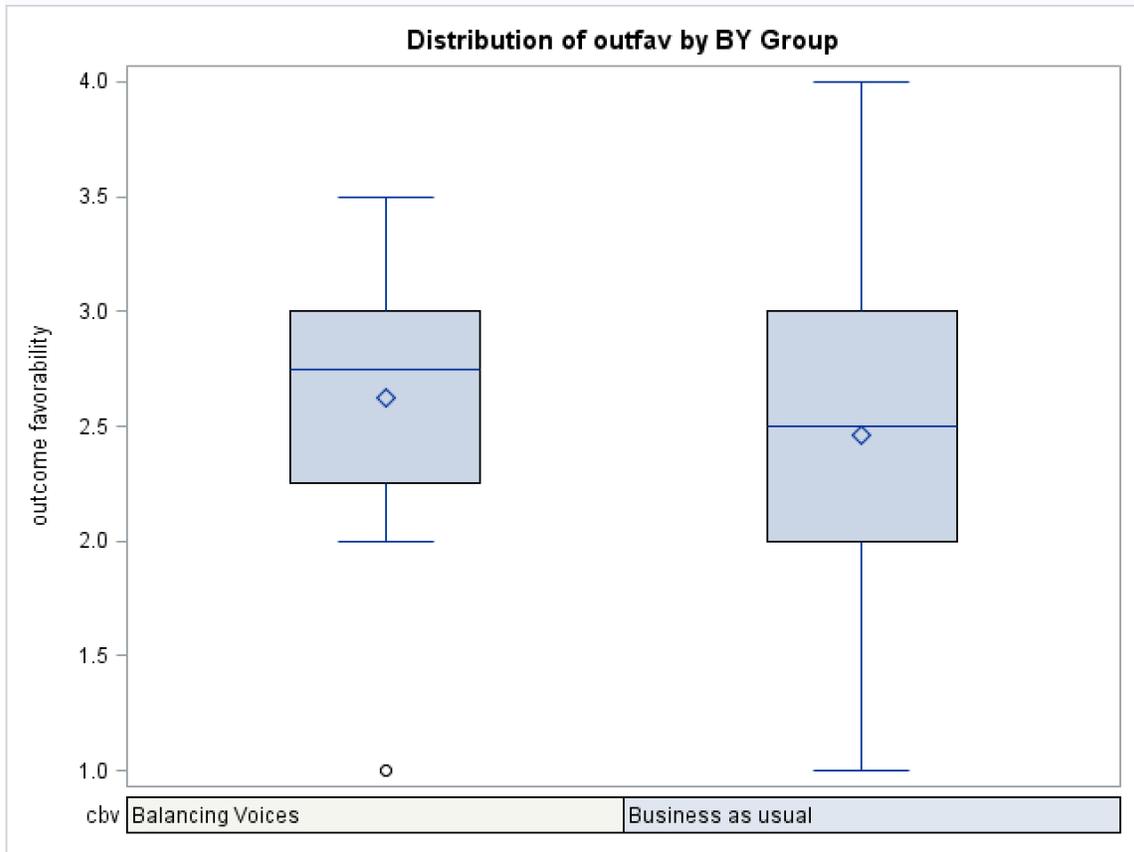


Figure 3b. Boxplots of Outcome Favorability by Protocol

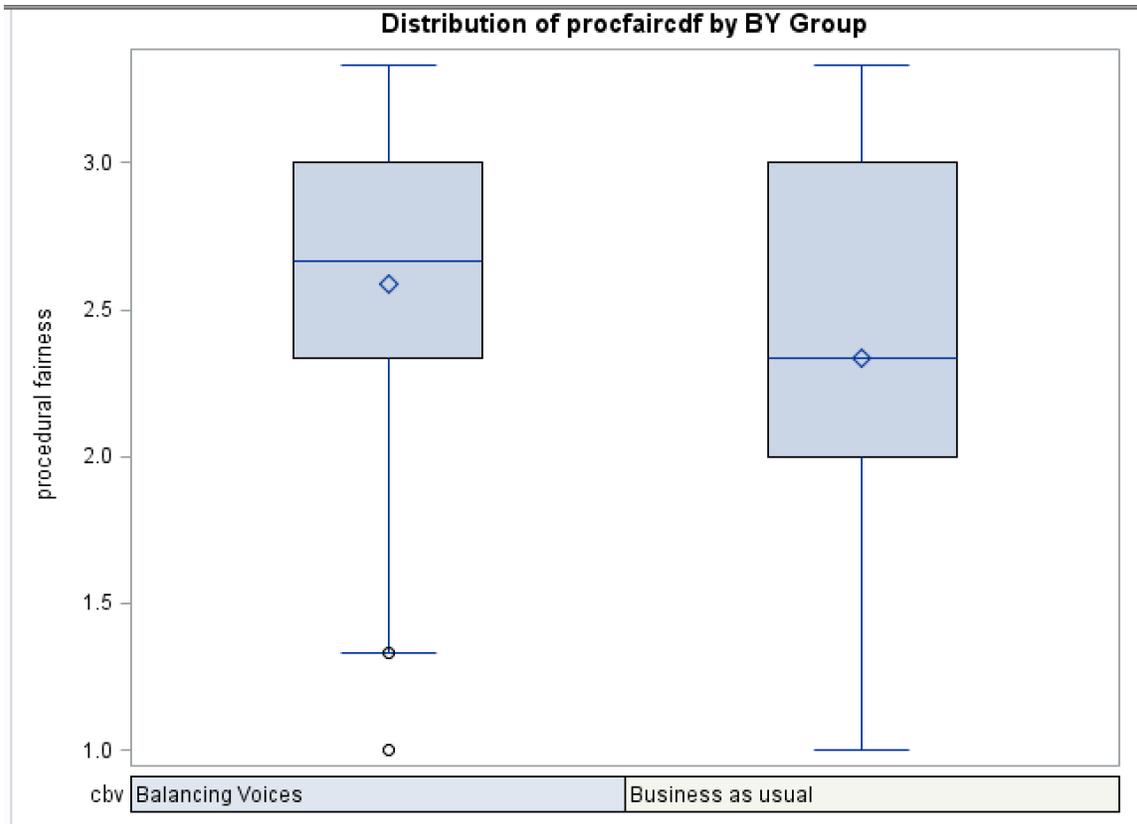


Figure 3c. Boxplots of Legitimacy of Procedural Fairness by Protocol

Table 1: Balancing Voices and Conventional Protocols for Decision-making in Schools

Goal	Action	Balancing Voices (Red Cards)	Business as Usual (Black Cards)
<i>1. Enhance educators' investment in, and commitment to, reforms to ensure success and sustainability</i>	Adoption of reform	No school-wide reform or change in policy or practice may be implemented unless the principal and two thirds or more of the teachers present to vote approve the change.	The state or district determines a school's curriculum
<i>2. Allow time for educators to balance fidelity of implementation of reform with local adaptation through collaboration</i>	Evaluation of reform	The effects of any change in practices or policies on student achievement should not be evaluated in terms of standardized test scores for three years after the initial vote to adopt the change. Short term effects can be evaluated in terms of changes in school climate (collective trust), teacher turnover, teacher absenteeism, student turnover.	The success of a new curriculum or reform is evaluated in terms of changes in test scores after it is implemented
<i>3. Coordinate community voice in school affairs</i>	Formal community bodies	The school shall have a local school council that has partial responsibility for governance. The council is to be composed of at least 50% of community members who are parents or guardians of students in the school. The council can vote to replace a principal by a vote of two thirds or more.	A community elects the members of a school board, which oversees all educational matters including the hiring and firing of administrators.
<i>4. Enhance formal voice for teachers concerning school administrators</i>	Evaluation and dismissal of a principal	A principal must be evaluated for replacement by the district if more than 20% of the teachers in the school (or 80% of the teachers in a given department) request transfer or leave in a given year (excluding planned retirements) or if 50% or more of the teachers sign a petition requesting the principal's removal.*	A district superintendent can fire or remove a principal at the end of a school year because the principal is ineffective or negligent
<i>5. Encourage ongoing teacher improvement and quality with due process</i>	Evaluation and dismissal of a teacher	A principal can use a streamlined procedure to remove not more than 5% of the teachers in a given year. The other 95% of teachers must be removed with full due process (3 years of documentation).	Due process (3 years of documentation) is required for a principal to remove a teacher

Table 2. Descriptive Statistics of Outcomes by Protocol

Variable	Balancing Voices	Conventional
	Mean (std)	Mean (std)
Procedural Fairness	2.59 (.55)	2.33 (.60)
Legitimacy of Authority	2.88 (.63)	2.48 (.74)
Figures		
Outcome favorability	2.63 (.59)	2.46 (.71)
N	61	62

Table 3: Estimated Effects of Balancing Voices on Outcome Favorability, Procedural Fairness and Legitimacy of Authority Figures

Outcome	No Controls	Control for Scenario and Role	Control for Participant Random Effects	All Controls
Procedural Fairness	.25* (.10)	.22* (.01)	.26* (.10)	.20* (.1)
Legitimacy of Authority Figures	.40** (.13)	.37** (.12)	.40** (.13)	.32** (.11)
Outcome favorability	.16(.13)	.13 (.11)	.16 (.13)	.11 (.11)

n=123

None of the participant random effects were statistically significant either by themselves or after adding the treatment predictor.

^ap < .10; * p < .05; ** p < .01; *** p < .001

Table 4: Interaction between Balancing Voices and Outcome Favorability

	Legitimacy of Authority		Procedural Fairness	
Outcome favorability	.32*** (.09)	.39*** (.11)	.38*** (.08)	.46*** (.09)
Balancing Voices	.34** (.12)	.78 (.47)	.19 (.09)*	.78* (.36)
Balancing Voices x outcome favorability		-.17 (.18)		-.23 ^a (.14)

Models include controls for scenario, role, and person random effects.

^ap < .10; *p < .05; **p < .01; ***p < .001