Understanding How Principal Evaluation Policies Inform Principal Perceptions and Focus on Learning Centered Leadership

Madeline Mavrogordato
Michigan State University

Morgaen Donaldson
University of Connecticut

Shaun Dougherty
University of Connecticut

Peter Youngs
University of Virginia

Introduction

Research indicates that principals play a crucial role in establishing conditions in schools that promote high-quality teaching and high levels of student learning. In particular, when principals engage in learning-centered leadership (LCL), they are more likely to foster instructional improvement and gains in student achievement (Bryk & Schneider, 2002; Grissom, Loeb, & Master, 2013; Newmann, Smith, Allensworth, & Bryk, 2001; Robinson, Lloyd, & Rowe, 2008). LCL involves (a) establishing rigorous goals for student learning, (b) coordinating curriculum, (c) enforcing high standards for teaching, (d) supporting teacher learning, (e) maintaining productive relationships among school staff, and (f) cultivating connections with families and community members (Bryk, Sebring, Allensworth, Luppescu, & Easton, 2010; Hallinger & Heck, 1996; Leithwood & Janzi, 2005; Murphy, Elliott, Goldring & Porter, 2007).

While there is strong evidence of the importance of LCL for student learning and school performance, it is less clear how school district policy can contribute to improved leadership practices among principals. In recent years, the great majority of states have implemented policies mandating that principals be evaluated each year based in part on student academic performance (Center on Education Policy [CEP], 2014; Jacques, Clifford, & Hornung, 2012). On one hand, recent changes in principal evaluation are based in policymakers' conviction that school leaders play a very important role. On the other hand, there has been little attention in the research literature to whether new district principal evaluation policies address leadership behaviors aligned with LCL either in their written form or when they are enacted in practice.

We address these questions by analyzing principal evaluation policies from 20 school districts in Connecticut and Michigan with respect to how they emphasize four broad leadership domains as identified by Goldring and colleagues (2009): instruction (e.g., monitors instruction creates climate of learning), management (e.g., oversees school facilities, follows fiscal policies),

personal traits (e.g., demonstrates written communication skills, applies ethical behavior), and community relations (e.g., promotes the school, engages with parents). Recognizing that intended policy, as represented by district documents, is not equivalent to the way policy is implemented in practice, we gauge the extent to which principals perceive that principal evaluation policies focus on these domains when enacted in practice. Specifically, we ask:

1. To what extent do districts' formal written principal evaluation policies emphasize instruction, management, personal traits, and community relations?

2. To what extent do principals perceive their district's principal evaluation policies emphasize instruction, management, personal traits, and community relations when enacted in practice?

3. Is there a relationship between the extent to which formal written principal evaluation policies emphasize the domains of instruction, management, personal traits, and community relations and principals' perceptions of how much these policies emphasize these domains when enacted in practice?

In the first section of this paper, we review previous research on principal evaluation and policy and policy content analysis. In the second section, we draw on institutional theory as a conceptual framework to examine the relationship between what written principal evaluation policies emphasize and principals' perceptions of policy focus in practice. In the third section, we present our methods including our data and sample, measures, and analytical strategies. The fourth section features our main findings regarding the extent to which district principal evaluation policies address the domains of instruction, management, personal traits, and community relations. Finally, we discuss implications of our findings for district efforts to promote effective leadership practices.

Research on District Principal Evaluation Policy

In the wake of the 2009 *Race to the Top* (RTTT) program and Title I waivers during the Obama administration, many states and districts enacted changes in principal evaluation. Recent studies of district principal evaluation systems have documented a primary emphasis on promoting principals' skills and a secondary focus on holding them accountable. But few studies have investigated ways in which district principal evaluation policies address different leadership domains, either as written or as enacted in practice. For example, in a study of changes in principal evaluation in six large urban districts, Anderson and Turnbull (2016) reported that new systems included measures of professional practice and student achievement growth and that all six districts used evaluation primarily as a tool to improve principals' leadership skills. In each district, the role of principal supervisor had changed to focus less on monitoring school leaders' compliance with district priorities and more on supporting their growth as instructional leaders. Finally, novice principals in each district received individualized support from mentors or coaches.

Examining principal evaluation systems in three large urban districts and two small rural districts, Kimball, Arrigoni, Clifford, Yoder, and Milanowski (2015) concluded that new systems were more rigorous and complex than earlier ones. Some districts in their study provided formal professional development for principal supervisors to ensure that they implemented the evaluation system with fidelity. In contrast, other districts depended on state training of principal supervisors and/or offered informal support for them. The authors further reported that most districts had enacted goal setting and continuous improvement cycles for principals, and the three large districts had reduced supervisor-principal ratios. Finally, a few districts supplemented principal evaluation with mentoring and coaching (Kimball et al., 2015). Neither Kimball et al.

(2015) nor Anderson and Turnbull (2016) focused on whether LCL was addressed by district principal evaluation policy.

Other research has examined the tensions district leaders encounter when enacting principal evaluation systems. For example, Zepeda, Lanoue, Price, and Jimenez (2014) examined principal evaluation practices in one district following enactment of RTTT and identified several tensions encountered by the district superintendent. First, there were sometimes discrepancies between ratings of a principal's performance and school performance. Second, there were persistent questions about how much emphasis to place on student achievement data versus other measures of leadership or performance. Third there was often a need to adjudicate between a principal's self-rating and the final ratings of their performance. Finally, the superintendent had strong knowledge of and relationships with each of the district's principals and had to be careful not to let these factors bias his ratings of them (Zepeda et al., 2014).

Lastly, Author (2015) investigated how principals in one district made sense of and responded to multiple sources of feedback on their leadership effectiveness. Principals interpreted feedback by making comparisons between their self-ratings and their teachers' ratings, between their teachers' ratings and teachers' ratings of other principals in the district, and between their current and past ratings. Consistent with research on feedback (e.g., Kluger & DeNisi, 1996) some principals adopted self-protecting behaviors and defensive orientations when faced with critical feedback, particularly in cases where they had rated themselves higher than their teachers had rated them (Author, 2015).

Taken together, these studies suggest that districts concentrated on developing principals' skills and holding them accountable in response to RTTT and Title I waivers. In addition, districts invested in training for principal supervisors and prioritized feedback and in some cases coaching for principals within the context of evaluation. At the same time, few of these studies

examined the types of leadership that were addressed by district principal evaluation policies either as written or as enacted in practice.

Policy Content Analysis

Scholars have analyzed written policies to determine their focus and change over time. They have analyzed the content of policies issued by the U.S. Department of Education on technology use in schools (Roumell & Salajan, 2016), policies aimed at reducing bullying in K-12 settings (Smith et al., 2012; Roberge, 2011), states' policies on special education inclusion (Duhaney, 1999), and equity-oriented reforms (Trujillo & Woulfin, 2014). Scholars use content analysis to identify specific portions of a policy that reflect particular aspects of the conceptual framework.

For example, in their analysis of the content of policies on technology use in schools, Roumell and Salajan (2016) drew on several frameworks describing policy implementation overall and specific to technology initiatives to identify nine "facets" of federal-level policy. These included "Policy provisions for requisite infrastructure—improving access and connectivity" and "Policy provisions for multiple funding streams and sustainability." The authors then coded each policy document line-by-line, assigning the nine facets to relevant excerpts from the policy. After engaging in various reliability procedures, the authors then quantified the degree to which each document reflected each of the nine facets. By applying these procedures to code a series of documents representing different points in time, Roumell and Salajan (2016) were able to draw conclusions regarding how the emphases of federal policy on K-12 technology use changed over time. Content analysis thus enables researchers to describe the specific substance of policies and quantify the extent to which various content appears in a policy.

Analyses of Principal Evaluation Policies

Few scholars have systematically analyzed written principal evaluation policies. Those who have engaged in this work have tended to analyze broad trends in principal evaluation. For example, Fuller and colleagues (2015) examined principal evaluation policies in all 50 states and Washington, DC. They reported that the main purpose of these state policies was to support school leaders' professional growth; in all, 78% of states articulated the improvement of principals' skills as an explicit purpose of evaluating them. In contrast to pre-RTTT findings, over two-thirds (68%) of states tied evaluation results to the compensation, promotion, or dismissal of school leaders. Consistent with federal policy, most state plans contained measures of student performance. The authors also found that in 27% of the states principal evaluation systems included data on teachers' working conditions/school climate and in 21% of states such systems featured measures of teacher quality, effectiveness, and/or retention (Fuller et al., 2015). In short, the measures included in principals' evaluations are more diverse and more tied to student and teacher outcomes than prior to RTTT.

While research such as Fuller et al.'s describes new evaluation systems in broad strokes, it does not shed light on the constructions of leadership advanced by new evaluation systems or how these systems are implemented. Most relevant for our study, Goldring et al. (2009) analyzed principal evaluation instruments in 68 urban school districts in 43 states. Goldring and colleagues analyzed and characterized the extent to which principal evaluation instruments reflected four domains: Instruction, Management, Personal Traits, and Community Relations. They found that few of the instruments addressed leadership behaviors – such as promoting rigorous curriculum or enforcing high standards for teaching – that are associated with student achievement (Goldring et al., 2009). Our study builds upon this work by investigating not only the extent to which our sample districts' written principal evaluation policies focus on these leadership

domains, but also whether principals perceive a similar emphasis when policies are implemented in practice. For example, if a written policy places emphasis on principals as instructional leaders, do principals perceive this emphasis when the policy is implemented in practice in their district?

Conceptual Framework

In this study, we draw on institutional theory as a conceptual framework in order to examine the association between the focus of principal evaluation policy documents and principals' perceptions of policy focus in enacted evaluation policies. Institutional theory provides tools for understanding associations between intended (i.e., written) educational policies and how they are actually implemented in practice; it emphasizes how professional norms, beliefs, and practices are constructed and reconstructed over time (Coburn, 2001). Carriers and institutional elements are two key concepts that institutional theorists employ to explain this process. Carriers are the actors, relationships, resources, and schema that transport new organizational forms between and within organizational structures (Anagnostopoulos et al., 2010). Examples of carriers in our study include district administrators and principals, relationships among them, and schemas held by principals about what principal evaluation is and should be.

Carriers employ three types of institutional elements to shape the potential adoption of new organizational forms: regulative, normative, and cognitive-cultural. Regulative elements promote change through formal policy, regulations, rewards, or sanctions (Anagnostopoulos et al., 2010; Scott, 2001). In our study, the regulative elements are represented by each state's principal evaluation policy and the ways in which each district modifies their written principal evaluation policies to be in line with state policy.

Normative elements potentially create new professional roles or change educators' internalized expectations and obligations (Anagnostopoulos et al., 2010; Scott, 2001). Examples of normative elements in our study include past and current practices in the area of principal evaluation, including such evaluation components as student achievement gains, principals' approaches to evaluating teachers, observations of leadership practices, and principal portfolios.

Cultural-cognitive elements potentially create new professional identities, categories of action, and perceptual schemas (Anagnostopoulos et al., 2010; Scott, 2001). In our study, cultural-cognitive elements include shared understandings of principal evaluation policies or how school leaders think about effective leadership broadly, as well as how it is measured. For example, some school leaders may define their role as maintaining strong community relations and attribute their success to personal traits even if the new principal evaluation policy in their district focuses on measuring their instructional leadership. Alternatively, it may be that most principals perceive an evaluation emphasis on instructional leadership regarding of the extent to which their district policy actually concentrates on such leadership.

Together, all of these different institutional elements shape whether new organizational forms take hold by promoting certain logics, or frameworks, through which school actors (i.e., district administrators, principals) make sense of and behave in accordance with within their environment (Scott, 2001). These logics are often deeply held and endure over time, but multiple logics typically co-exist within fields and they are continually contested and negotiated (Anagnostopoulos et al., 2010). It is therefore possible that logics can be altered or replaced with competing logics, such as a policy change like the one under exploration in our study. The potential conflict between and among institutional elements and logics is an area of notable interest in our study because it allows us to more deeply examine similarities and differences

between how district principal evaluation policies as written and how they are enacted in practice (i.e., perceived by principals).

In this study, we draw on institutional theory to compare written district principal evaluation policies with how principals perceive that they are actually implemented in practice. Institutional theory illuminates the processes through which principal evaluation reform interacts with school leaders' occupational beliefs and structures by specifying the regulative, normative, and cultural-cognitive elements embedded in district policies. In particular, institutional theory allows us to examine whether, for example, policies that emphasize instructional leadership or management, as written, are perceived by principals to focus on these aspects of leadership as implemented.

Methods

We seek to understand the relationship between the types of leadership focused on by written principal evaluation policies and the emphases perceived by principals. Accordingly, we analyzed written policy documents and principals' perceptions.

Sample

We purposively selected districts from Connecticut and Michigan, which vary by socio-political context, average student achievement, and policy context. Both of these states grant districts discretion in developing their own principal evaluation systems, but have state policy that set specific evaluation parameters. Connecticut's state principal evaluation policy mandates that a principal's overall rating is composed of 45% student learning; 40% observation of leadership practice; 10% stakeholder feedback; and 5% teacher effectiveness. Michigan's policy dictates that at least 25% of a principal's rating is based on student achievement growth and assessment data and "the portion of the evaluation that is not based on student growth data and the district's adopted evaluation tool must be based on the administrator's proficiency in using

the observation tool for teachers; the progress made by the school or district in meeting the goals set forth in the school or district improvement plan as applicable; student attendance in the school or school district; and student, parent, and teacher feedback" (Michigan Department of Education, 2016).

Within these states, we selected districts to maximize variation in student enrollment, urbanicity, student demographics and district principal evaluation policies. Specifically, 10 participating districts were in Connecticut and 11 were in Michigan, for a total of 21 school districts. Districts ranged in enrollment from 1,000 to 10,000 students. More information about sample districts is provided in Table 1.

[Insert Table 1 here]

Data Sources

Data sources include principal evaluation documents and surveys of principals in each participating district. For each district, we collected district policy documents related to principal evaluation during the 2016-17 academic year. These included the written evaluation policy, evaluation rubrics, and corresponding documentation. For this paper, we focused specifically on rubrics to understand the types of leadership emphasized in principal evaluation policies across the districts in the sample.

Our second source of data came from surveys of principals in each participating district. In 2016-17 we administered a survey to principals working in schools across the sample districts. This survey inquired about principals' experiences with principal evaluation, including the degree to which they perceive the evaluation process to focus on instructional leadership activities versus managerial leadership activities; whether their performance is regularly monitored; whether principal evaluation results are used to plan principal professional development; and whether such results are linked to principal rewards and sanctions. The survey

was developed by the researchers based on prior work (Hamilton et al., 2012; Marsh et al., 2011; Sun & Youngs, 2009). The survey also included items on principals' demographic characteristics (e.g., race/ethnicity, gender) and professional experience (e.g., years of experience as principal, years of teaching experience). Of the 99 principals surveyed, 62 responded to the survey for an overall response rate of 63%.

Measures

  Written evaluation policy focus. Following Goldring and colleagues' (2009) content analysis of principal evaluation policies, we used an iterative, deductive process to code principal evaluation rubrics on each of four domains: instruction, management, personal traits, and community relations. We coded each district's evaluation rubric at the indicator level (e.g., each row of a rubric) using these categories and 36 sub-categories (coding scheme available upon request). Specifically, we read each indicator and, based on its content, tagged it with the appropriate sub-categories, which aggregate up to the four overarching domains.

  We coded in three stages. First, two members of the team selected a district's evaluation tool and each individual coded the tool independently. After completing the independent coding assignment, the two members met to discuss the general trends in the coding and to work through the coding of the rubric indicator by indicator. Early comparison revealed a high degree of agreement between the coders. At this point, the pair identified any emergent discrepancies and discussed these differences at length. Typically, these discussions resulted in a consensus. In rare cases, however, consensus could not be reached and the two coders brought the issue to the larger research group for discussion and resolution. We followed this pattern of coding independently, meeting together for consensus coding, and bringing any unresolved issues to the larger group to work through each of the evaluation tools used by the districts in our sample.

We constructed two measures of the degree of emphasis on the four domains from each of the district principal evaluation policies. First, following Goldring et al.'s (2009) procedures, we constructed a measure representing the percentage of codes within a particular domain divided by the total number of codes applied to a district's rubric. This first approach implicitly defines the four main areas of focus as additive to a single whole, and so all measures provide an indication of relative emphasis on one domain compared to the other three. In the second approach we constructed a measure representing the percentage of rubric indicators coded with a particular domain. Indicators could be (and were) coded with multiple domains when their substance indicated it. As a result, this second measure captures the degree of focus on each domain relative to the theoretical maximum that any one area could be represented with respect to itself, but not with respect to the other three main areas of emphasis.

Evaluation policy focus in practice. Using principals' survey responses, we constructed measures of principals' interpretation of their evaluation policy's focus on the four domains of instruction, management, personal traits, and community relations. For each domain, we generated two measures. The first measure was constructed from a series of questions (between 3 and 15 questions for each domain) on the principal survey that measured a single, unified construct for each domain. In all cases, the questions appeared to measure a single, unified construct for each domain (Cronbach's α as follows: instruction [0.95], management [0.93], personal traits [0.93], and community relations [0.84]). For the second measure, we used one item that asked the respondents the degree to which they thought their district's policy focused on each of the four domains. For control variables, we also constructed standard measures to capture principal demographic information from the survey.

Analysis

Our key question predictors for Research Question #1 (*WrittenPolicyFocus*) is our researcher-coded measure of the extent to which a district's policy focuses on each of four leadership domains. To answer our primary research question, we fit the following statistical model, which relates principals' interpretation of the evaluation policy's instructional focus (*PerceivedInstructionFocus*) to our policy predictor (*WrittenPolicyFocus*) that was generated from coding the policy documents:

$$PerceivedInstructionFocus = WrittenPolicyFocus + X\_p + J\_s + Gamma\_state + epsilon$$

Though in the baseline specification we include only these two key variables, we also added principal characteristics (dummy variables for female and White, as well as years of experience as a principal), and a state fixed effect (to differentiate systematic differences across states). We applied an adjustment to account for deviations from the OLS assumption of homoscedastic standard errors, which were clustered at the school district level. We tested the sensitivity of this relationship to our multiple approaches to generating the *WrittenPolicyFocus* variable from our coding.

Analytic Strategy

Summary statistics for focal outcomes and the two different approaches to coding the policy focus, our principal predictors of interest, appear in Table 2. Overall, principals have a high level of understanding that their evaluation policies emphasize instruction. In fact, there may be a ceiling effect such that variation in this measure is reduced therefore attenuating relationships that include this measure. Predictors generated from the coding of district principal evaluation policies show a similar degree of focus on instruction with instruction showing up most frequently regardless of the approach and management and personal traits getting the next highest degrees of representation.

[Insert Table 2 here]

To answer our research questions, we fit two sets of models to estimate the relationship between principal understanding of the focus of their evaluation policies, and our researcher coded approaches to understanding written policy emphasis, where instruction, management, personal traits, and community relations are the four large areas of emphasis. In the first set of models, we used principal's perceptions of the instructional focus of their evaluation policies as the only outcome. In the second set of models, we expanded to include principal's understanding of the extent to which their policies emphasized management, personal traits, and community relations. Finally, we also fit a series of simple bivariate relationships across these four areas of interest given the modest size of our dataset and related concerns about statistical power.

Results

Overall, in our hypothesized models, we find no statistically or practically significant associations between our researcher coded measures of written policy focus and principal's own reported perceptions of the emphasis of those policies on instruction, management, personal characteristics, or community relations. However, it is difficult to establish the relationships of interest in our current data in part because our dataset is small and all variation in written policy focus is at the district level. Such associations also appear to be difficult to establish because there is limited variation in the measures, even among our small dataset. In more simplified bivariate models we find positive and statistically significant associations between measure of management focus and personal traits. Importantly, in the models that associate management measures, we find that adding in the coded measure of instructional focus adds important insights.

Estimates in Table 3 follow the general structure of our hypothesized statistical model and use a measure of principals' perception of the extent to which their evaluation policy focuses

on instruction as the outcome. In columns 1-4 we used our first method for coding the policy documents, and columns 5-8 we used a second coding method to produce these measures. In columns 1-3 and 5-7 we use the multiple question measure of instructional focus, whereas in columns 4 and 8, we used the response to a single question as the outcome. In columns 4 and 8 we also include as predictors measures of the extent to which policies were coded as focused on management and personal traits. Columns that include controls for gender, race (binary indicator for white) and years of experience as a principal are indicated as including these controls. All models include an indicator for the state, to account for differences that might be attributable to state, and not district, policy requirements.

[Insert Table 3 here]

We present the results of fitting a second set of models in Table 4. In these models we used composite measures of a principal's understanding of the extent to which their evaluation policies focus on instruction, management, personal traits, and community relations as the outcomes. These outcomes are composites from the related series of questions in the principal survey which are noted in the column headings. In columns 1-4 we use independent variables generated using coding method 1, and columns 5-8 use measures generated from the second method.

[Insert Table 4 here]

Considering our estimates from Table 3, we do not see any statistically significant association between the outcome, principals' perceptions of instructional policy focus, and the primary predictor of interest, instructional focus of the written policy. However, it is worth noting that across models, the relationship between the outcome and instructional focus is suggestively negative, except in the presence of the other policy focus variables, at which point the sign and magnitude of the relationship change. Also noteworthy is that in models that include

multiple measures of policy focus, there are always larger point estimates (relative to instruction or personal characteristics) associated with management focus, regardless of the model or the version of the policy coding.

Results presented in Table 4 suggest that when modeling a principal's understanding of their policy's focus on instruction, management, or personal traits, the independent variable management focus has the strongest relationship (in terms of magnitude), relative to the other predictors of interest, though all relationships are imprecise (one relationship is statistically significant, though the number of models we fit suggest this could be type 1 error). When using community engagement as the outcome management focus loads as negative, though the strength and direction of the relationship (even suggestive) for the other predictors, depends one which approach to policy coding is used.

To further understand the lack of association in our statistical models, we present a series of Figures (1 through 4) that demonstrate the association between principal's perception of the extent to which their evaluation policy emphasizes a given domain, and the coded levels of that factor in the principal evaluation policy document. In Figure 1, we graph the relationship between a principal's perception of the importance of instruction in their principal evaluation system (y axis) against the standardized (mean zero, variance 1) measure of emphasis of the district policy document on instruction. In this Figure, we label districts by state where a 1 indicates a district in Connecticut, and a 0 indicates a Michigan district. In Figures 2 through 4 we repeat the same process using our respective measures of management, personal traits, and community relations.

[Insert Figures 1-4 here]

Figure 1 makes clear that there is no evident relationship between a district's policy focus on instruction and a principal's perception of the importance of instruction in their policy. In

16

fact, the uniformly high levels of principal perception of the importance of instruction seems to suggest that this understanding is independent of the extent to which their written evaluation policies emphasize it.

In contrast to our first figure, Figures 2 and 3 suggest that there may be some positive association between principal's perceptions of the emphasis of management and personal traits in their evaluation system and the extent to which the coded written policies appear to emphasize these factors. However, as with instructional focus, Figure 4 (and associated regression results) does not show any indication of an association between principals' perception of the importance of community relations and the extent to which the coded policy emphasizes this dimension.

As suggested by this graphical analysis, we also fit a series of associated regression models, the results of which we present in Table 5. Columns 1, 2, 3, and 5 correspond to Figures 1, 2, 3, and 4, respectively. The lack of a statistically significant relationship between the instruction and community relations measures is evident from the estimates in columns 1 and 5, whereas the suggestive positive linear association among management and personal trait measures is clear in columns 2 and 3.

[Insert Table 5 here]

Given the positive statistically associations in columns 2 and 3, we also added instructional focus to those models and report those results in columns 4 and 6. In column 6 we find that adding instructional focus into the model that looks at personal traits measures does not attenuate the relationship between personal traits measures, and that instructional focus is positively associated with principal's perceptions of the focus on personal traits, though this association is not statistically significant.

Our most interesting finding comes from adding instructional focus to our model associating the management measures. This addition does not undermine the positive significant

association between the coded measure of management, and principal's perceived importance of management. In fact, instructional focus is also significantly, though negatively, associated with principal's perceptions of the importance of management. This suggests that, for any given level of management emphasis in a policy, a greater degree of emphasis on instruction will tend, on average, for principals to perceive that management is less important. This finding comports with our general hypothesis, based on the literature, that on average school leaders perceive a tradeoff between management and instructional leadership responsibilities.

Discussion

In sum, we found no relationship between the substantive focus of written principal evaluation policies, based on policy coding, and principals' reports regarding the emphasis of different domains in these policies, controlling for key covariates. In bivariate analyses, however, we found a positive association between written policy focus on personal traits and principals' perceptions that this domain was emphasized. We also found a positive relationship between policy focus on management and principals' perceptions of a management emphasis. We must treat these findings with caution, however, given the small sample size and preliminary nature of these analyses.

If these findings persist as we expand our sample and refine our analyses, they raise several important observations and questions. First, the most reasonable conclusion we might reach at this stage is that principals have a general perception that instructional focus is of high import regardless of the relative weight instructional focus appears to receive in a variety of different approaches to coding these policies. Lack of relationship between instructional focus based on policy coding and instructional focus, reported by principals, is on the one hand not surprising. Rhetoric advocating instructional leadership is commonplace today. Preparation programs emphasize this sort of leadership and our interviews with superintendents of districts in

this sample suggest that they believe that instructional improvement should be principals' priority (Donaldson, Mavrogordato, Youngs & Dougherty, 2017). The lack of relationship between a written policy's instructional focus and principals' perceptions of instructional focus could be because principals are not aware of the details of the policy and and are receiving (or reporting) strong messages that their policy focuses on instruction. First, central office may be ignoring the district's actual policy and overemphasizing instructional leadership, relative to the actual policy, in their communications with principals. Second, central office may be sending signals aligned with their policy's relative emphasis on the four domains, but principals are interpreting these signals as placing a emphasis on instructional leadership that dwarfs the weight central office administrators and the policy itself asign to it. Third, principals may be aware of the relative emphasis of instructional leadership in the policy but are responding to social desirability in reporting an outsized focus on instructional leadership. All of these potential explainations could be at play, and we intend to explore these hypotheses as we inteview sample principals in 2018.

Once we include predictors or outcomes that include management or personal traits, we find suggestive evidence that written policy focus on management and personal traits is positively related to principals' understanding of the emphasis on each of these respective domains. If these findings hold, it appears that the policy messages that stress the importance of management (e.g., managing school facilities) and personal traits (e.g., maintaining integrity, responsibility) somehow get through to principals. It could be that these messages are distinct enough from the others that central office administrators convey information about these domains more clearly and that principals take note of them. Interviews with district superintendents suggested that they viewed management and personal traits as first-order concerns; they reported that when evaluating principals they did not venture into instructional

19

domain if a principal struggled to keep her building orderly or failed to follow through, take responsibility for her actions, or behave ethically (Donaldson et al., 2017). Some districts may encounter these first-order concerns more frequently and consider them more salient than their district counterparts do, leading them to construct principal evaluation policies that emphasize management or personal traits and send signals to principals to focus on these types of leadership. In turn, principals in these districts may be more likely to report that their evaluation systems focus on management or personal traits. We will investigate these possibilities as we expand our sample and collect a new round of data.

Lastly, we find an interesting relationship between management focus in the written policy and principals' perceptions of the policy's management focus, controlling for instructional leadership focus in the written policy. Specifically, we find that the relationship between management focus in written policy and principals' perceptions remains strong and positive, but that the instructional focus of the written policy is marginally significant and negatively related to principals' perceptions of a management focus in the policy. In other words, the less the policy actually focuses on instructional leadership, the more that principals report a policy emphasis on management. This suggests that principals may perceive instructional leadership and management as trade-offs, that they channel their energy towards management goals at the expense of instructional leadership and vice versa. This finding, if it holds with our expanded sample, adds evidence to a set of studies that hint at complex interplay between managerial and instructional leadership (e.g., Grissom, Loeb & Master, 2013). This finding also reflects an longstanding and ongoing tension between the two dominant conceptions of principal leadership among practitioners, instructional leadership and management (e.g., Bossert, Dwyer, Rowan & Lee, 1982; Cuban, 1988; Murphy et al., 2007).

References

Anagnostopoulos, D., Sykes, G., McCrory, R., Cannata, M., & Frank, K. (2010). Dollars,

distinction or duty: The meaning of the National Board for Professional Teaching Standards

for teachers' work and collegial relations. *American Journal of Education, 116*(3), 337-369.

Anderson, L.M. & Turnbull, B.J. (2016). *Evaluating and supporting principals.* New York, NY:

Wallace Foundation.

Author. (2015).

Author. (2017).

Bossert, S. T., Dwyer, D. C., Rowan, B., & Lee, G. V. (1982). The instructional management

role of the principal. *Educational Administration Quarterly*, *18*(3), 34-64.

Bryk, A.S., & Schneider, B. (2002). *Trust in schools: A core resource for improvement*. New

York, NY: Russell Sage Foundation.

Bryk, A.S., Sebring, P.B., Allensworth, E., Luppescu, S., & Easton, J.Q. (2010). *Organizing

schools for improvement: Lessons from Chicago*. Chicago: University of Chicago Press.

Center on Education Policy. (2014). Federal education programs: NCLB/ESEA Waivers.

Washington, DC: Author. Retrieved from http://www.cepdc.org/index.cfm?Document

SubTopicID=48.

Coburn, C. (2001). Collective sensemaking about reading: How teachers mediate reading policy

in their professional communities. *Educational Evaluation and Policy Analysis, 23*(2), 145-

170.

Cuban, L. (1988). *The managerial imperative and the practice of leadership in schools*. Albany,

NY: Suny Press.

Donaldson, M., Mavrogordato, M., Dougherty, S. & Youngs, P. (2017, November). *Principal*

evaluation policies on paper vs. in practice: Evidence from 20 districts. Paper presented at the annual meeting of the University Council for Educational Administration, Denver, CO.

Duhaney, L.G. (1999). A content analysis of state education agencies' policies/position statements on inclusion. *Remedial And Special Education*, *20*(6-), 367-78.

Fuller, E.J., Hollingworth, L., & Liu, J. (2015). Evaluating state principal evaluation plans across the United States. *Journal of Research on Leadership Education, 10*(3), 164-192.

Goldring, E., Cravens, X.C., Murphy, J., Porter, A.C., Elliott, S.N., & Carson, B. (2009a). The evaluation of principals: What and how do states and urban districts assess leadership? *Elementary School Journal, 110*(1), 19-39.

Grissom, J.A., Loeb, S., & Master, B. (2013). Effective instructional time use for school leaders: Longitudinal evidence from observations of principals. *Educational Researcher, 42*(8), 433-444.

Hallinger, P. & Heck, R.H. (1996). Reassessing the principal's role in school effectiveness: A review of empirical research, 1980-1995. *Educational Administration Quarterly, 32*(1), 5-44.

Hamilton, L.S., Engberg, J., Steiner, E.D., Nelson, C.A., & Yuan, K. (2012). *Improving school leadership through support, evaluation, and incentives: The Pittsburgh Principal Incentive Program.* Santa Monica, CA: RAND.

Jacques, C., Clifford, M. & Hornung, K. (2012). State policies on principal evaluation: Trends in a changing landscape. Washington DC: National Comprehensive Center for Teacher Quality.

Kimball, S.M., Arrigoni, J., Clifford, M., Yoder, M., & Milanowski, A. (2015). *District leadership for effective principal evaluation and support.* Washington, DC: U.S. Department of Education, Teacher Incentive Fund.

Kluger, A. & DeNisi, A. (1996). The effects of feedback interventions on performance: A historical review, a meta-analysis, and a preliminary feedback intervention theory. *Psychological Bulletin*, *119*(2), 254-284.

Leithwood, K., & Janzi, D. (2005). A review of transformational school leadership research, *Leadership and Policy in Schools, 4*(3), 177-199.

Marsh, J.A., Springer, M.G., McCaffrey, D.F., Yuan, K., Epstein, S., Koppich, J., Kalra, N., DiMartino, C., & Peng, A. (2011). *A big apple for educators: New York City's experiment with schoolwide performance bonuses. Final Evaluation Report.* Santa Monica, CA: RAND.

Murphy, J., Elliott, S. N., Goldring, E., & Porter, A. C. (2007). Leadership for learning: a research-based model and taxonomy of behaviors. *School Leadership and Management*, *27*(2), 179–201.

Newmann, F.M., Smith, B., Allensworth, E., & Bryk, A.S. (2001). Instructional program coherence: What it is and why it should guide school improvement policy. *Educational Evaluation and Policy Analysis, 23*(4)*, 297-321.*

Roberge, G.D. (2011). Countering school bullying: An analysis of policy content in Ontario and Saskatchewan. *International Journal Of Education Policy And Leadership*, *6*(5), 1-14.

Robinson, V.M.J., Lloyd, C.A., & Rowe, K.J. (2008). The impact of leadership on student outcomes: An analysis of the differential effects of leadership types. *Educational Administration Quarterly, 44*(5), 635-674.

Roumell, E.A., & Salajan, F.D. (2016). The evolution of U.S. e-learning policy: A content analysis of the National Education Technology Plans. *Educational Policy*, *30*(2), 365-397.

Scott, W. R. (2001). *Institutions and organizations* (2nd ed.)*.* Thousand Oaks, CA: Sage.

Smith, P.K., Kupferberg, A., Mora-Merchan, J.A., Samara, M., Bosley, S., & Osborn, R. (2012).
A content analysis of school anti-bullying policies: A follow-up after six years. *Educational Psychology In Practice*, *28*(1), 47-70.

Sun, M., & Youngs, P. (2009). How does district principal evaluation affect learning centered principal leadership? Evidence from Michigan school districts. *Leadership and Policy in Schools, 8*(4), 411-445.

Trujillo, T. M., & Woulfin, S. L. (2014). Equity-oriented reform amid standards-based accountability: A qualitative comparative analysis of an intermediary's instructional practices. *American Educational Research Journal*, *51*(2), 253-293.

Zepeda, S.J., Lanoue, P.D., Price, N.F., & Jimenez, A.M. (2014). Principal evaluation – linking individual and building-level progress: Making the connections and embracing the tensions. *School Leadership & Management, 34*(4), 324-351.

Table 1

Information on Sample Districts

District Pseudonym	State	Enrollment	% Economically Disadvantaged	% Students of Color	% English Learner	% Special Education	Urbanicity
Bradley	CT	21,000	100%	90%	15%	15%	Midsize City
Washington	CT	7,000	65%	85%	10%	15%	Small City
Elmer	CT	5,200	35%	20%	<5%	15%	Large Suburb
Oakwood	CT	2,000	10%	10%	<5%	10%	Large Suburb
Gaffney	CT	3400	10%	15%	<5%	10%	Large Suburb
Mayville	CT	1,200	25%	25%	5%	15%	Rural-Fringe
Morrison	CT	6,200	20%	20%	<5%	15%	Small City
Spaulding	CT	2,400	10%	10%	<5%	10%	Rural-Fringe
Valliant	CT	3,300	45%	30%	5%	20%	Large Suburb
Warner	CT	3,100	35%	65%	5%	20%	Large Suburb
Carleton	MI	4,700	65%	50%	10%	15%	Small City
Norwood	MI	1,100	55%	20%	10%	10%	Rural-Distant
Clearmont	MI	2,700	50%	10%	15%	5%	Distant Town
Lambert	MI	3,000	15%	10%	<5%	5%	Large Suburb
Rhine	MI	3,400	25%	5%	<5%	10%	Large Suburb
Barrett	MI	8,400	30%	20%	5%	10%	Large Suburb
Gorman	MI	5,700	40%	25%	5%	10%	Large Suburb
Hamilton	MI	2,000	50%	25%	10%	10%	Distant Town
Jefferson	MI	6,400	5%	50%	15%	10%	Small City
Ralston	MI	5,200	30%	10%	5%	15%	Large Suburb
Parkston	MI	2,200	40%	10%	<5%	10%	Town-Fringe

Note: Numbers are rounded to protect the anonymity of participating districts. Data on economically disadvantaged students in CT retrieved from Kids Count Data Center 2013-2014, kidscount.org. Data on economically disadvantaged students in MI retrieved from CEPI 2014-2015 data, mischooldata.org. All other data collected from the Common Core of Data 2014-2015 LEA survey.

Table 2
Summary Statistics for Variables

Variable	Obs	Mean	Std. Dev.	Min	Max
Outcomes: Principals' Perceived Policy Focus					
Instruction	62	3.36	0.56	1.90	4.00
Management	62	2.50	0.85	1.00	5.00
Personal traits	62	2.44	0.70	0.88	3.63
Community relations	62	3.00	0.74	1.00	4.33
Written Policy Focus Coding Measure 1					
Instruction	62	0.66	0.09	0.55	0.80
Management	62	0.07	0.03	0.03	0.17
Personal traits	62	0.14	0.05	0.06	0.27
Community relations	62	0.12	0.07	0.02	0.22
Written Policy Focus Coding Measure 2					
Instruction	62	0.83	0.08	0.70	1.00
Management	62	0.17	0.07	0.08	0.35
Personal	62	0.32	0.14	0.13	0.52
Community relations	62	0.27	0.17	0.05	0.72

Table 3

Relationship Between Written Policy Emphasis and Perceived Instructional Emphasis

	Policy Code Method 1				Policy Code Method 2			
	(1) B5 Composite Perceived Instruction	(2) B5 Composite Perceived Instruction	(3) B5 Composite Perceived Instruction	(4) B1 Single item Perceived Instruction	(5) B5 Composite Perceived Instruction	(6) B5 Composite Perceived Instruction	(7) B5 Composite Perceived Instruction	(8) B1 Single item Perceived Instruction
Written Policy Emphasis								
Instruction	-0.0740	-0.275	0.173	0.180	0.0670	-0.0550	0.358	-0.212
	(1.017)	(1.128)	(1.959)	(1.663)	(0.916)	(0.962)	(1.155)	(1.186)
Management			2.484	1.462			0.979	1.215
			(2.946)	(2.926)			(1.205)	(1.237)
Personal			0.673	-0.747			0.197	-0.544
			(2.613)	(2.431)			(0.733)	(0.753)
Connecticut	0.061	0.072	0.071	0.166	0.0713	0.100	0.0499	0.104
	(0.188)	(0.192)	(0.237)	(0.228)	(0.148)	(0.150)	(0.162)	(0.166)
Controls	N	Y	Y	N	N	Y	N	N
Intercept	3.381***	3.645***	3.127	3.599*	3.271***	3.525***	2.810*	3.890**
	(0.727)	(0.751)	(1.650)	(1.531)	(0.781)	(0.804)	(1.107)	(1.137)
N	62	62	62	62	62	62	62	62

Standard errors in parentheses

* p<0.05, ** p<0.01, *** p<0.001

Table 4

Relationship Between Written Policy Emphasis and Perceived Emphasis in Four Leadership Domains

	Policy Code Method 1				Policy Code Method 2			
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
	B5 Composite Perceived Instruction	B2 Composite Perceived Management	B4 Composite Perceived Personal Traits	B3 Composite Perceived Comm Relat	B5 Composite Perceived Instruction	B2 Composite Perceived Management	B4 Composite Perceived Personal	B3 Composite Perceived Comm Rel
Written Policy Emphasis								
Instruction	0.173	1.588	-0.811	-0.822	0.150	-0.0410	1.738	1.522
	(1.959)	(2.812)	(2.294)	(2.560)	(1.219)	(1.711)	(1.451)	(1.579)
Management	2.484	7.372	7.941*	-1.361	1.195	2.589	2.368	-1.017
	(2.946)	(4.229)	(3.450)	(3.850)	(1.228)	(1.724)	(1.461)	(1.591)
Personal	0.673	6.838	2.873	0.172	0.0372	1.589	1.195	1.326
	(2.613)	(3.751)	(3.061)	(3.415)	(0.742)	(1.041)	(0.883)	(0.961)
Connecticut	0.0710	0.129	-0.206	-0.139	0.0628	0.0862	0.0209	0.0230
	(0.237)	(0.340)	(0.277)	(0.309)	(0.163)	(0.229)	(0.194)	(0.211)
Female	0.0213	0.0646	0.0881	0.0781	0.0503	0.172	0.186	0.0744
	(0.151)	(0.217)	(0.177)	(0.198)	(0.151)	(0.211)	(0.179)	(0.195)
White	-0.362	-0.129	-0.399	-0.631	-0.395	-0.225	-0.532	-0.641
	(0.259)	(0.372)	(0.304)	(0.339)	(0.258)	(0.362)	(0.307)	(0.334)
Total Years Principal	0.0378	-0.179	-0.0163	-0.125	0.0296	-0.172	-0.0946	-0.170
	(0.0843)	(0.121)	(0.0987)	(0.110)	(0.0738)	(0.104)	(0.0878)	(0.0956)
Intercept	3.127	0.534	2.435	4.618*	3.231**	2.235	0.911	2.569
	(1.650)	(2.369)	(1.933)	(2.157)	(1.135)	(1.593)	(1.351)	(1.471)
N	62	62	62	62	62	62	62	62

Standard errors in parentheses
* $p<0.05$, ** $p<0.01$, *** $p<0.001$
Notes: Community focus predictor is omitted because other three measures total determine the focus of the policy, by construction

29

Table 5

Relationship Between Written Policy Emphasis and Perceived Emphasis in Four Leadership Domains

	Policy Code Method 1					
	(1)	(2)	(3)	(4)	(5)	(6)
	B5 Composite Perceived Instruction	B3 Composite Perceived Community	B2 Composite Perceived Management	B2 Composite Perceived Management	B4 Composite Perceived Personal traits	B4 Composite Perceived Personal traits
Written Policy Emphasis						
Instruction	-0.012			-2.195+		1.478
	(0.896)			(1.293)		(1.315)
Community Relations		0.391				
		(0.578)				
Management			3.565*	3.192*		
			(1.488)	(1.481)		
Personal					1.143+	1.627*
					(0.641)	(0.771)
Intercept	3.367*	2.895*	1.892*	3.785*	2.081*	0.695
	(0.750)	(0.181)	(0.272)	(1.147)	(0.221)	(1.252)
N	62	62	62	62	62	62

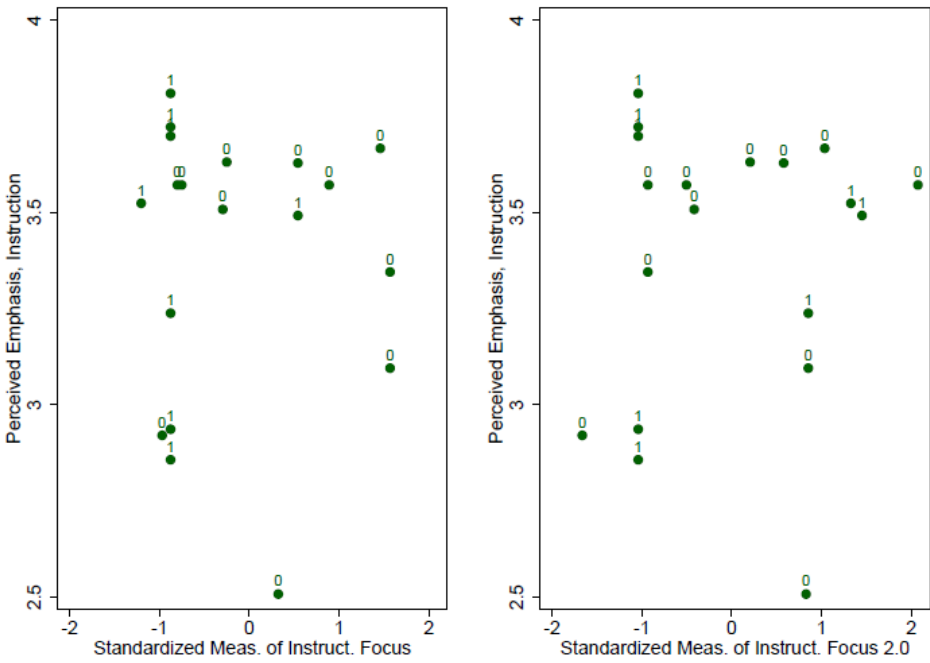Standard errors in parentheses

+ $p<0.10$, * $p<0.05$

Figure 1. Relationship between principal understanding of instructional focus (Y) and policy coded measure of instructional focus (X) using two different approaches to policy coding.
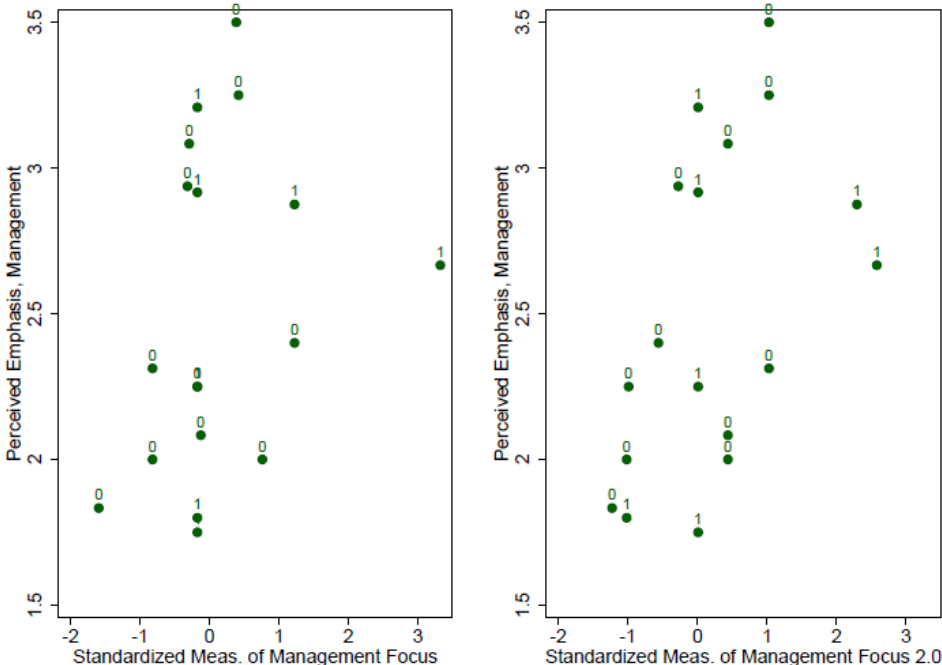

Figure 2. Relationship between principal understanding of management focus (Y) and associated policy coded measure (X) using two different approaches to policy coding.
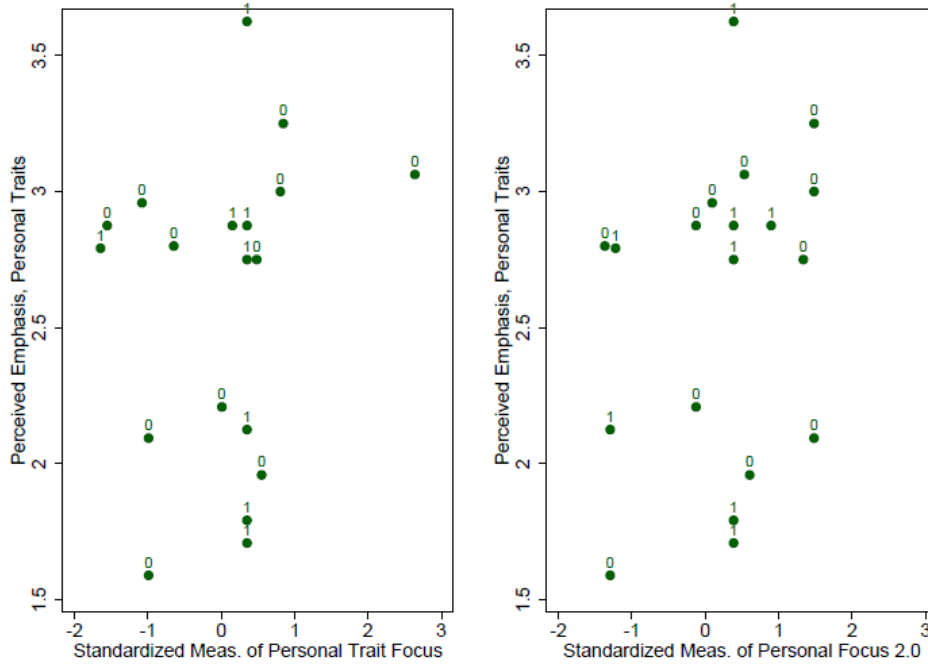
Figure 3. Relationship between principal understanding of personal traits focus (Y) and associated policy coded measure (X) using two different approaches to policy coding.
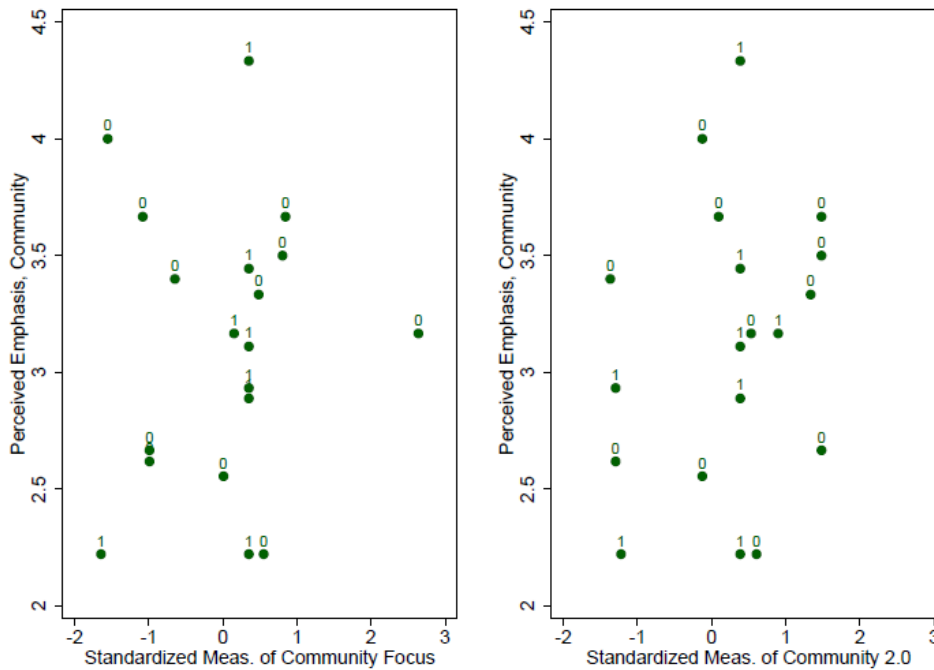

Figure 4. Relationship between principal understanding of community relations focus (Y) and associated policy coded measure (X) using two different approaches to policy coding.