The Quality of Leaders' Problem-Solving Conversations:

Truth-Seeking or Truth-Claiming?

Viviane Robinson, Frauke Meyer, Deidre Le Fevre, Claire Sinnema

The University of Auckland

Author details:

Distinguished Professor Viviane Robinson

The University of Auckland

Private Bag 92601, Symonds St, Auckland 1150,

T. + 64 6238899 x 87372 E-mail: vmj.robinson@auckland.ac.nz

Dr Frauke Meyer

The University of Auckland

Private Bag 92601, Symonds St, Auckland 1150,

T. + 64 6238899 x 48471, E-mail: f.meyer@auckland.ac.nz
Orcid: 0000-0002-7402-8896

Dr Deidre Le Fevre

The University of Auckland

Private Bag 92601, Symonds St, Auckland 1150,

T. + 64 6238899 x 89843, E-mail: d.lefevre@auckland.ac.nz

Associate Professor Claire Sinnema

The University of Auckland

Private Bag 92601, Symonds St, Auckland 1150,

T. + 64 6238899 x 46426, E-mail: c.sinnema@auckland.ac.nz

The Quality of Leaders' Problem-Solving Conversations:

Truth-Seeking or Truth-Claiming?

Abstract

Effective leaders solve problems which prevent the achievement of organizational goals. We drew on theories of problem-solving and interpersonal behavior to propose a process model of quality problem-solving conversations, which was used to evaluate three leaders' attempts to resolve problems in conversation with those involved. Analyses revealed problematic patterns: little disclosure of causal beliefs, little public testing of beliefs that might trigger negative emotions, and agreement on solutions misaligned to causal beliefs. This research suggests that leaders need development opportunities that help them to frame their beliefs in ways that avoid creating a dilemma between maintaining relationships and solving problems.

Although considerable attention has been given by educational researchers to the role of goal-setting in leadership effectiveness (Meyer, Sinnema, & Patuawa, 2019, Sinnema, & Robinson, 2012; Sun & Leithwood, 2015), there is far less known about the ability of leaders to overcome the numerous problems that stand in the way of goal achievement (Mumford, Todd, Higgs, & McIntosh, 2017; Mumford, Zaccaro, Harding, Jacobs, & Fleishman, 2000). Without adequate problem-solving capability, leaders will struggle to formulate and resolve problems in ways that create the goal achievement and improvement they seek (Saiti, 2015).

The quality of leaders' problem-solving matters because the decisions they make about what constitutes a problem and how it should be addressed have important ethical and educational consequences (Dempster & Berry, 2003). For example, resources may be wasted, and students' learning stymied, because school leaders introduce new programs and innovations without a careful prior diagnosis of the learning needs of the target students. Whether such 'quick fixes' are a response to external political pressure (Bryk, Gomez, Grunow & Le Mahieu, 2015) or to school leaders' own drive for improvement, this pattern of problem-solving often brings costs in teacher burnout and cynicism without the intended benefit of improved learning and teaching (Meyer, Sinnema & Patuawa, 2019). Given that so much is at stake, there is a moral and professional obligation on leaders to be diligent about how problems are identified, understood and resolved.

But what counts as diligence? What does high-quality problem-solving look like? What does research tell us about the level of leaders' capability or, more importantly, about the match between their capability and that required by the problems they seek to resolve? In this paper we explore these questions by investigating the quality of leaders' problem-solving as they talk about their on–the-job problems with those who are directly involved. There are several reasons why, despite the methodological and ethical challenges of such research, we have chosen to study leaders' problem-solving through analysis of their real conversations. First, the problems that educational leaders are responsible for are inherently *social* in the sense that their resolution requires the coordinated action

3

of individuals. Investigating leaders' capability in social problem-solving requires close study of how they think, feel and talk during meetings and conversations with those whom they are attempting to influence. This is not possible when leaders make individual responses to problems presented as written scenarios, which is the methodology that has been used in much previous research (Allison & Allison, 2003; Brenninkmeyer & Spillane, 2008; Leithwood & Stager, 1989; Spillane, White, & Stephan, 2009). While use of scenarios offers the considerable advantage of standardized problems, its downside is that it strips away the emotional and interpersonal dimensions that are at play when leaders tackle problems that cannot be solved without influencing their colleagues. Scenarios are unlikely to evoke the emotional challenges that proved, in this study, to be such a powerful driver of leaders' problem-solving.

A second reason for contextualizing problem-solving within real conversations is that this context provides data about leaders' actual problem-solving practice rather than their self-reported practice. When leaders report or reflect on how they did or would solve a problem, in the absence of behavioral data on their actual practice, we cannot judge the match between their espoused theory of problem-solving and their actual theory-in-use (Argyris, 1976). In this study conversation transcripts provided the behavioral data.

In sum, our research approach provides an authentic context for studying how educational leaders attempt to solve problems that are real, that involve their colleagues and that are emotionally and cognitively taxing. In studying problem-solving in this way, we are responding to Mumford et al.'s (2017) call for "far more research, […] on the skills leaders need to solve the complex problems, inherently social problems, they are expected to solve" (p.13).

Two Theories of Problem-solving

The description and evaluation of leaders' problem-solving capability is critically dependent on the normative theory of problem-solving employed. Two such theories have dominated the literature. The first, more rational model proposes a linear sequence of problem identification,

evidence-based analysis of causes, formulation of solution strategies that are tightly linked to causes and to intended outcomes, followed by solution implementation and evaluation (Bransford & Stein, 1993; Simon, 1993). While there are many variations of this rational model, it prescribes reduction of uncertainty through careful processing of information and strong alignment between hypothesized causes and proposed solutions.

A major criticism of this rational model is that it is mismatched to the speed, complexity and uncertainty of leaders' work (Robinson, 2014; Mintrop & Zumpe, 2019). An alternative, intuitive theory of problem-solving suggests that, rather than understanding problems through a careful search and analysis of relevant information, leaders take a more holistic approach to ill-structured problems – that is to problems in which a) goals are contested and dilemmatic, b) there are multiple possible paths to their achievement, and c) what counts as success is unclear (Simon, 1993). The intuitive theory suggests that as leaders gain different types of experience, they develop schema and prototypes through which they recognize situations as being of a particular type and, therefore, as susceptible to a familiar solution (Sternberg & Horvarth, 1995). In other words, their cognitive schema enables them to make a quick match between the situation and what they have already learned about how to deal with it. While this 'first to mind' strategy enables swift action, untested assumptions and various cognitive biases mean that it may not lead to a quality solution (Kahneman, 2011). The question remains, therefore, of how to formulate a normative theory that fosters problem-solving quality while addressing the complexity, uncertainty and fast pace of the work (Milkman, Chugh & Bazerman, 2009).

Research on the Problem-solving of Educational Leaders

In a series of studies of educational leaders (Leithwood & Stager, 1989; Leithwood & Steinbach, 1995), Leithwood and his colleagues developed a multi-component model of problem-solving that blends aspects of both theories of problem-solving. Two of the six components are concerned with problem formulation (interpretation and goal setting), two with actions taken to

solve the problem (constraints and solution processes), and two with underpinning values (principles/values) and emotion (affect). The studies compared the problem-solving practices of expert leaders (as judged by school performance indicators and leader reputation), with those of non-expert leaders, using written scenarios in one and analysis of real problem-solving meetings in the other (Leithwood & Steinbach, 1995). In general, experts did more investigation of the problem before offering solutions, situated problems in a wider strategic context, made more detailed action plans and monitored and evaluated the implementation of solutions more thoroughly than their non-expert counterparts. These findings suggest that experts adopt a more deliberate, systematic and rational approach to problem-solving than non-experts.

Other scenario-based research suggests that the expert non-expert distinction may not always be clear cut. In a quantitative comparison of the problem-solving strategies of principals classified as expert (n=20) or non-expert (n=16) in school improvement, Brenninkmeyer and Spillane (2008) report mostly non-significant differences between the two groups on 19 problem-solving processes, which were mainly derived from the measures used in Leithwood's research program. The differences, however, were largely in the expected direction and varied depending on the content of the six different instructional leadership scenarios employed. In a very similar study, Spillane, White and Stephan (2009) used the same scenarios to test for differences between expert and aspiring principals on 22 problem-solving processes. Significant differences were found for only 5 of the 22 processes and, for two of those processes, the differences were contrary to those hypothesized. Expert principals were more likely than aspiring principals to make assumptions about the problem, possibly reflective of their greater experience with the type of problem, and to give less consideration to the long-term strategic implications of their problem-solving.

A quite different methodology was used by Mintrop and Zumpe (2019) to study how nine expert leaders enrolled in a doctoral program framed the problems they encountered in their school-based improvement projects. Rather than employing scenarios and the problem-solving indicators

used by Leithwood and Spillane, they took a more grounded theory approach by using documents, assignments and structured interviews to track, over two years, how these expert school leaders interpreted and attempted to resolve their chosen problems. Their predominant cognitive strategy was to swiftly categorize their problems as being amenable to a particular solution. Indeed, it was the absence of their preferred solution that largely framed their interpretation of the problem - a strategy that is consistent with the intuitive rather than rational theory of problem-solving. Even when prompted by instructors and coaches to check their assumptions by gathering relevant data, the inquiry stance of these expert leaders was directed towards confirmation rather than disconfirmation of their assumptions. The issue with their problem-solving was not that they sometimes proposed solutions before inquiring into causes, but that they did not recognize that their solutions presumed the validity of their fallible beliefs about the problem's likely causes.

In summary, it appears that a clear-cut distinction between the problem-solving of expert and non-expert educational leaders may not be defensible, with variations in subject matter, the mode of problem representation and the domain specific knowledge of problem-solvers shaping their use of more rational or more intuitive approaches. Given what is at stake, however, we are still left with the need for a contextually sensitive normative model of problem-solving. On the one hand, swift intuitive processes reduce cognitive load and enable faster decision-making (Kahneman, 2011). On the other hand, the possibility of bias and error, with its substantial repercussions for the lives and learning of teachers and students, is an ever-present possibility. The main contribution of this paper is to propose a normative model for quality problem-solving that recognizes that either approach, or a mix of the two, may be appropriate in particular contexts. For us, quality is indicated by the extent to which, *at any stage* of a problem-solving conversation, leaders check rather than assume the validity of their key beliefs about the nature and resolution of their important on the job problems. This means, for example, that our concern is not to identify whether a solution is proposed at the outset or after careful inquiry into the nature and cause of a problem. Rather it is whether the likely validity of the solution is checked at any stage during the conversation.

Validity and Validity Testing

The concept of validity employed in this paper reflects an epistemological stance in which truth is treated as a regulative ideal – one that is never finally attained but worth striving for if we value learning and quality problem-solving (Phillips, 1987). Thus, leaders who enact this ideal treat their beliefs as fallible, and their reasoning as potentially flawed, while, nevertheless, striving to improve them through a process of error detection and correction (Argyris, 1976; Mazutis & Slawinski, 2008). In short, their stance is one of truth-seeking rather than truth-claiming.

Beliefs may be assumed be true, that is not tested at all, or their validity can be privately or publicly tested (Argyris & Schön, 1974; Stanovich & West, 1997). When leaders publicly test their beliefs, they disclose them along with the grounds they are based on, so their accuracy or reasonableness can be evaluated by those who may hold a different interpretive frame (Dewulf & Bouwen, 2012). When leaders test their beliefs privately, they do so without disclosing them and involving others in checking their validity. Instead, they use information about the object of their belief to make private judgments about whether their view is justified. Mistaken assumptions and questionable inferences are less likely to be detected in private than in public testing, because the same logic that generated the belief is used to evaluate its validity (Argyris, 1976). If this logic remains unquestioned only single-loop solutions are possible – i.e. those that are compatible with the leader's initial assumptions about the problem (Argyris, 1976; Robinson, 2014). Public testing opens up the possibility of double-loop learning i.e. of learning from others that what was taken-for-granted about the problem is inaccurate or incomplete (Argyris, 1976; Robinson, 2014).

The idea of public testing has a long history in ethics (Hare, 2003; Spiegel, 2012), philosophy of science (Popper, 1999) and interpersonal and organizational learning (Argyris 1990; Argyris & Schön, 1996). These authors make clear that public testing, in the context of problem-solving conversations, involves the enactment of particular interpersonal and epistemic values, rather than the mere use of particular behaviors. As Spiegel (2012) writes, leaders must be willing to reconsider

their views, and such reconsideration is "not a matter of mere perfunctory listening to contrary opinions but a genuine readiness to revise or even abandon one's views in light of new objections or counter evidence" (p.8). On this view, empirical investigation of leaders' validity testing should access leaders' cognitions as well as behaviors (Mumford, Watts, & Partlow, 2015).

Validity testing in problem-solving conversations

In this section, we proposed a process model of quality problem-solving in which quality is indicated by the extent to which leaders publicly test the validity of their beliefs about the nature, causes of, and solutions to their problems (Figure 1).

[Insert Figure 1 here]

In outlining our model, we explain what we mean by beliefs, describe three values that underpin belief testing, and outline six behaviors involved in such testing. While the limited research available suggests that the outcomes of our proposed process model of problem-solving are likely to include progress in both problem resolution and increased relational trust (Mumford, Gibson, Giorgini & Mecca, 2014; Tjosvold, Sun, & Wan, 2005; Zand, 2016), testing those outcomes was not part of our study.

Belief types

Beliefs express a link between an object (self, other, institution, event, or situation) and one or more attributes (Fishbein & Azjean, 1975). A leader might believe, for example, that a colleague's performance (the object) is unsatisfactory (the attribute). Beliefs are distinguished from knowledge by the fact that the latter and not the former have to satisfy a truth condition. While a belief may be "accepted as true by the individual holding the belief" (Richardson, 1996, p.104), the strength of the belief does not determine its epistemic warrant.

Problem-solvers bring hundreds of beliefs to any conversation and it is cognitively and practically impossible to test the validity of them all. Our proposed model recognizes the need for fast and efficient problem solving by focusing on the extent to which leaders tested their key beliefs

i.e. those for which validity is likely to be critical to successful problem resolution. Three types of belief were treated by the researchers as key. Since testing whether others agree that there is a problem is an important initial step in recruiting others to the problem-solving process (Mumford et al., 2000; Nickles, 1981), beliefs that describe a situation as problematic were treated as the first type of key belief and were called Problem Description Beliefs (PDB). The second type, Problem Explanation Beliefs (PEB) comprised beliefs about how the problem could be explained. Leaders' explanations were considered key because the probability of resolving a problem partly depends on finding solutions that address its likely causes (Mumford et al., 2017). Beliefs about how the problem could or should be solved constituted the third type of key belief – Problem Solution Beliefs (PSB). They were considered key, because such beliefs shape the action plans that are made, and the way resources are used (Mumford, Gibson, Giorgini, & Mecca, 2014).

Values supportive of public validity testing

Values are enduring beliefs about what is desirable in human conduct (Leithwood, Steinbach & Raun, 1993). In their explicit form, they are manifest in statements about what is important, and act as standards against which one's own and others' conduct is evaluated. In their implicit form, values are enacted in actual behavior. They act as automatic filters that rule in conversational moves as more or less acceptable in a given context. For example, a leader's value of "being positive" may rule in affirming and rule out critical feedback. In enacting values, tensions between them may be experienced as dilemmas or conflicts. For example, the value of "being positive" may be experienced as a constraint on other values such as being honest or empowering of others.

The first and central value of our proposed model is the desire to maximize the validity of the information used to solve a problem. Leaders who enact this value are open-minded and treat their views as fallible and thus subject to critique and revision (Spiegel, 2012). Their interpersonal stance enables "progressively more effective testing of assumptions and progressively greater learning about one's effectiveness" (Argyris & Schön, 1974, p. 86). This value is akin to accuracy

motivation (Hart, Albarracin, Eagly, Brechan, Lindberg, & Merrill, 2009) and can be contrasted with a closed-minded desire to win and to be right (Argyris & Schön, 1974).

One cannot be open to learning from others' views without the second value of respect. Respect for others is demonstrated when those involved are treated as having a right to express their own beliefs, as capable of learning, and of contributing to others' learning. It is disrespectful to withhold relevant views and to ignore or discount difference. Validity and respect are interdependent for it is the robust and respectful contest of ideas that provides tough tests of validity (Tjosvold, 2008). Without respect, leaders will not be able to build the relational trust needed to access and learn from others' ideas and to give and receive honest feedback about the validity of beliefs (Zand, 2016).

The third value involves increasing internal rather than external commitment to the problem-solving process and outcomes. Leaders who value internal commitment want others to follow agreed courses of action because they believe they are the right ones in the circumstances, not because they feel compelled to do so. Internal commitment is increased when doubts and disagreements are disclosed and discussed rather than withheld (Bambacas & Patrickson, 2008). It is also increased by the provision of choice about courses of action, after due public consideration of the validity of the information that is guiding the choice. In contrast, a closed-minded persuasive stance is more likely to lead to external rather than internal commitment to decisions.

Behavioral indicators of public validity testing

Figure 1 includes six behavioral indicators of public validity testing that are aligned to these three values. The first indicator, *disclosing beliefs*, is a precondition of public testing for without disclosure the other party can only speculate about what the leader really thinks. Disclosure aligns with the value of maximizing valid information motivated by truth-seeking and interpersonal respect.

The second indicator is *providing the grounds* on which the key belief is based. Grounds for beliefs may comprise reasons, examples, or relevant evidence. Leaders who aim to maximize valid information disclose their beliefs, explain how they reached them, and help others to do the same. When reasoning is explicated and evidence provided, others can independently assess and critique the likely validity of the beliefs, thereby increasing the chance of detecting and correcting faulty reasoning (Argyris & Schön, 1974, Colquitt, Scott, & LePine, 2007).

Inquiring into others' thinking about the problem constitutes the third validity testing behavior. Such inquiry can take many different forms, including asking directly for others' beliefs or for reactions to one's own. Less direct forms of inquiry include creating space for others' views by expressing one's own views in invitational language (Le Fevre & Robinson, 2015). Whatever the linguistic form of the inquiry, it contributes to public validity testing if it is motivated by a desire to learn from others and revise one's own views accordingly (Hare, 2003; Spiegel, 2012).

The fourth validity testing behavior, *engaging with difference*, captures whether leaders inquire into or discuss the implications of differences or disagreement for the validity of their beliefs (Dewulf & Bowen, 2012). Such engagement contrasts with the more frequent reactions to difference which are restatement or elaboration of one's own position, and rebuttal of the other's beliefs (Le Fevre & Robinson, 2015).

Checking the alignment between beliefs about the nature (PDB), causes of (PEB) and solutions to (PSB) the problem is our fifth validity testing behavior. As we have already discussed, the resolution of complex problems does not require that discussion of possible causes precedes discussion of possible solutions. Regardless of the cause-effect sequence apparent in a problem-solving conversation, our view is that, given the high stakes involved in most educational problem-solving, explicit discussion of the alignment between problem explanations and proposed solutions is central to quality problem-solving (Zaccaro et al., 2000). Failure to inquire into this alignment

can lead to repeated cycles of single-loop learning, producing temporary "fixes" rather than a lasting solution (Argyris, 1976; Bokeno, 2003).

Reaching agreement about a key belief is the sixth validity testing behavior. The other party's expression of agreement ought to be explicit, that is, directly stated and freely given rather than coerced. It must be acknowledged, however, that agreement is not a wholly reliable indicator of validity as, at times, it may not be warranted (Phillips, 1987). In a problem-solving conversation, however, leaders must strive to gain a level of agreement that motivates pursuit of the problem-solving process. We can detect possible unwarranted agreement, however, by checking whether it was reached in combination with any of the other five validity testing behaviors.

Given the paucity of empirical data about the frequency and consequences of these six behaviors, we can only speculate about their relative importance. As we discussed earlier, however, it can be argued, on theoretical grounds, that engaging with difference is a particularly powerful validity testing behavior as it deliberately exposes key beliefs to the possibility of disconfirmation. In contrast, the 'reaching agreement' behavior provides a weaker test of validity because agreement may be reached without any such exposure. The complexity, uncertainty and high stakes nature of many educational problems, would suggest that it is combinations of these validity testing behaviors, rather than single occurrences, that characterizes quality problem-solving.

Research Questions

Three research questions guided the investigation of different aspects of the proposed model. The first question, "Which key beliefs are leaders more or less likely to disclose?" recognizes that disclosure of beliefs is a prerequisite for public testing and that there may be selective disclosure of different types of belief. Answers to our second question "How do leaders test the validity of their disclosed key beliefs?" will tell us about leaders' differential use of the five remaining validity testing behaviors. In answering the third question, "Which values do leaders enact in their problem-solving conversations?" we probed the relationship between leaders' validity

13

testing behaviors and their underlying values, as revealed by their unexpressed thoughts and feelings during the conversation. We addressed this question by an analysis of the enactment of the three values in our model, and by a more inductive search for evidence of other values that could explain the degree of validity testing.

Method

In this section, we outline the research context and participants before describing the data collection and analytic strategies we used to answer our three research questions.

Research Context

Our data were gathered from part-time students enrolled in a graduate course in educational leadership. The students, who held management roles in schools, gathered evidence about their own leadership behavior by recording a conversation with a colleague from the educational setting in which they were employed. The purpose of the conversation was to discuss a concern they held about their colleague's performance or behavior. The word "concern" was treated as synonymous with "problem" and used for data collection purposes because earlier studies revealed leaders' reluctance to use the term "problem" in conversation with colleagues (Sinnema et al., 2013).

While students had read some initial articles by Argyris and Schön before they recorded their conversation, they had not rehearsed or discussed their conversation beforehand. They were urged to use their typical approach in the conversation so they could get an authentic sample of their on-the-job problem-solving behavior. It is possible, however, that the course readings had cued students to the importance of testing the validity of their beliefs.

Participant Sampling

Twenty-seven students completed the recording and consented to its use for research purposes. Prior to examining the transcripts of the recordings, the authors established criteria for drawing a subsample from the 27 available conversations for the detailed qualitative analysis reported in this study. Inclusion in the subsample required that the leader: held a senior educational leadership

position (principal, deputy, associate or assistant principal); had at least four years' experience as a senior leader, and that they had made prior attempts to resolve their concern. This latter criterion increased the probability of investigating possible links between leader behavior and problem persistence. Only three of the 27 leaders met all three criteria and they constitute the participants in this study. All three leaders were white, had graduate qualifications, and were either assistant or deputy principals in urban elementary or high schools. All three had more than ten years teaching experience. Pseudonyms are used for the names of all individuals and schools.

We report our in-depth analysis of data from just three participants, to contribute precise explanatory insights about the phenomenon of interest (validity testing in the talk of educational leaders). We recognize that a larger sample would be required if our purpose was to establish generalizability of our claims. Our purpose is different—to understand through highly detailed analyses how what leaders think and say impacts on the extent to which their conversations serve to increase or decrease the validity of the beliefs that are bought to bear on problem-solving efforts.

Such detailed analyses allow us to attend to what Ball (2018) refers to as "discretionary spaces". In these discretionary spaces, micro-level practices occur that shape and are shaped by macro-level educational issues. In this study, the macro-level educational issues which were relevant to our leaders' conversations included the difficulty of school improvement, the persistence of performance problems and, ultimately, inequity in student achievement. Our sample of three allowed rigorous attention to the ways in which these more macro-level issues played out, at the micro-level, in leaders' interactions with their colleagues.

An intensive focus on three cases enabled us to integrate data from multiple sources and establish a clear audit trail from our original data sources to the inferences we made about the validity testing behavior of each leader (Morse & Field, 1995; O'Reilly & Parker, 2012). It also enabled us to carefully compare the quality of problem-solving across the three belief types and the three leaders, and to provide enough contextual information to make sense of the comparisons.

Data Sources

Prior to holding their conversation, leaders completed a short questionnaire (Appendix A), comprising nine open-ended questions and five rating scales, about the nature and history of their problem and their prior attempts to resolve it. Evidence about how leaders thought, felt and acted while engaged in problem-solving was obtained from an annotated transcript of an unrehearsed conversation with the person who was the subject of their concern. Leaders audio-recorded a six to eight-minute conversation which they then transcribed. The conversation transcript was entered in the right-hand column (RHC) of a split page, along with line numbers. In the left-hand column (LHC) leaders were asked to annotate the transcript with what they were thinking and feeling during the conversation but did not share at the time. Leaders were instructed not to include evaluative reflections on their action in their annotations, but to retrieve, as far as possible, their reflections in action (Schön, 1991). Argyris and Schön (1974), who first developed this procedure, refer to these annotations as unexpressed thoughts and feelings or the left-hand column. Recording both speech and associated unexpressed thoughts and feelings, enabled investigation of the relationship between the more surface features of validity testing, as evidenced in leaders' actual speech, and their enacted values, as evidenced in their speech and their undisclosed thoughts (Hart et al., 2009). While this LHC procedure provides an important window into leaders' interpersonal values, it is important to acknowledge that leaders may have differed in their willingness to disclose their unexpressed thoughts and feelings.

Data analysis

The analysis strategy involved three steps: identifying leaders' key beliefs, classifying the type of belief and then determining whether and how leaders tested the validity of each key belief. The twenty-page codebook used for all three steps was developed through two iterations of coding of three sets of transcripts and questionnaires drawn from the 24 leaders not included in our sample. It includes detailed rules and procedures for each analysis step.

16

The codebook development was an iterative process, involving independent coding by three of the four authors followed by discussion and resolution of differences. Resolution was achieved by discussion and clarification of the application of the theoretical model, and revision of the coding rules until they were interpreted and applied consistently (DeCuir-Gunby, Marshall, & McCulloch, 2011; MacQueen, McLellan-Lemal, Bartholow, & Milstein, 2008).

Identifying and classifying key beliefs

A key belief was defined as one whose validity was critical to successful problem resolution. Such beliefs, as expressed by the leader, were listed along with the data source(s) from which they were extracted. The sources were the questionnaire (Q) and the left-hand (LH) and right hand (RH) columns of the transcript. Once repetition had been eliminated, each belief was categorized into one of the three belief types explained earlier – beliefs about what the problem was (PDB); beliefs about how to explain it (PEB) and beliefs about how to solve it (PSB). The codebook included the definition, examples and coding rules for classification into a belief type. The coding rules guided the extraction and classification of key beliefs from the pre-conversation questionnaire as well as from the transcript. In the questionnaire, PDB beliefs were identified from leaders' problem descriptions (Item 1); PEB beliefs from questions about their own and others' contribution to the problem (Items 5-7) and PSB beliefs from a question about their views on what was needed to resolve the concern (Item 8). Occasionally, additional key beliefs were reported in answers to other questionnaire items and those beliefs were also coded.

The key belief and data sources for each belief were noted in the first column of the analysis tables (see Tables 1-3). One author took primary responsibility for identifying and classifying a leader's beliefs and a second author for checking the completeness and accuracy of both of these steps. The final list of key beliefs was confirmed after discussion of each author's coding and resolution of any differences between them.

Identifying patterns of validity testing

The first research question required a comparison of the types of key belief that were and were not disclosed. Disclosed beliefs were identified from the RHC of the leader's transcript. Undisclosed beliefs were those found in the questionnaire, or in the unexpressed thoughts and feelings (LHC) of the leader, but not in the leader's actual speech (RHC). When a belief was not disclosed no entries were made in the analysis table under the public validity testing behaviors. By comparing the three data sources across the three belief types, we could determine if particular types of belief were more or less likely to be disclosed.

The second research question about how leaders tested the validity of their key beliefs was answered by identifying from the transcripts the use of one or more of the six behavioral indicators of validity testing. The codebook for validity testing behaviors included rule descriptors, specific inclusion and exclusion criteria, and examples. First, a brief description of any use of a validity testing behavior was recorded, along with the transcript line reference, against the relevant key belief. While the focus of this description was primarily the leader's behavior, it was also noted whether the other party disagreed with a key belief of the leader, as such differences provided opportunities to test validity by discussing disagreement. This initial identification of validity testing behaviors was independently checked by two of the authors and any disagreements resolved by discussion of the relevant transcript excerpts. Once agreement was reached, the use of each validity testing behavior was recorded in the analysis table alongside the belief to which it applied

In answering the third question, "Which values do leaders enact in their problem-solving conversations?" we systematically compared, using the comprehensive display in each leader's table, the content of the beliefs which were tested with those that were not. This comparison of belief content provided clues about the reasons why the leader may have kept certain beliefs private. For example, a leader's disclosure of critical beliefs expressed by third parties, and non-disclosure of very similar beliefs held by the leader, suggests an interpersonal value of conflict avoidance or

protection of self from any negative emotion. All such inferences from the transcript about the leader's enacted values were cross checked against the annotations the leader made in the LH column of their transcript. These annotations usually included meta-level commentaries about the conversation itself, and this provided some insight into the extent to which leaders enacted the interpersonal values identified as important for collaborative problem-solving. For example, the left-hand annotation "I wonder if she agrees with me, I need to ask her" suggests respect for and openness to the other's view. A more closed-minded stance that is likely to inhibit validity testing is suggested by the comment "Maybe I've come on a bit strong – I'll try and bring him back around". This annotation indicates a desire, to keep the other person on side while trying another way of persuading them to one's own point of view. In reporting our findings, quotes from leaders' unexpressed (LHC) and expressed (RHC) thoughts and feelings are used to discuss relationships between their validity testing behaviors and their enacted interpersonal and epistemic values.

Findings: Three Leaders' Problem-Solving Approaches

The results of our analysis of the leaders' conversations are presented in the form of comprehensive tables (Tables 1-3) which list the content of each leader's key beliefs and their associated validity testing strategies. After a brief introduction to the context of the first leader's conversation (Caroline), we structure our discussion of the evidence in her table around our research questions to show which beliefs she disclosed, how she tested them, if at all, and what values she enacts. Quotes from her questionnaire and transcript are used throughout to illustrate the patterns revealed by the table. This analytic strategy is repeated for the second (Liz) and third (Paul) leader.

Caroline: Selective Disclosure and Testing Based on the Avoidance of Negative Emotion

Caroline, a deputy principal at an elementary school, chose to have a conversation with a teacher (Cheryl) whose loud voice and poor classroom management had been a cause of concern two years prior to this conversation. Some limited support, including observations, modelling and

feedback had been provided by the principal and Caroline, who was her team leader at the time. Cheryl's performance, however, did not improve enough to prevent a new principal moving her out of her classroom teacher role into a permanent teacher release position. Prior to the conversation, Caroline, who was now the deputy principal, wrote about how her concerns about Cheryl had resurfaced…

> Recently we had the Education Review Office [the national inspection agency] in the school, followed a few weeks later by math facilitators, who both commented on this teacher's 'loud' voice. The principal …asked me to meet with her, after the meeting he'd had with her a few days prior, to sort out what she was going to do about her voice and classroom management. He had asked me to observe the meeting as deputy principal and [union] representative. (Q, Item 1)

Despite the persistence of the problem, Caroline rated her prior attempts to solve it as 'satisfactorily' effective (Q, Item 16) and 'somewhat' influential (Q, Item 17). She now viewed the problem as 'very important' to resolve (Q, Item 18), but was only 'somewhat' confident in her ability to do so (Q, Item 19).

Which key beliefs does Caroline disclose?

In her conversation with Cheryl, Caroline disclosed two of her six beliefs about the nature of the problem, none of her three beliefs about its cause, and three of her six beliefs about how it could be solved (Table 1, VT1). Examination of belief content shows that she disclosed her principal's perceptions of the problem (PDB 1-2) and withheld her own (PDB 3-6), including her belief that some of the principal's concerns were valid (PDB 3). She did not disclose her beliefs about how the school (PEB 3) and Cheryl herself (PEB 1-2) had contributed to these long-standing difficulties, and also withheld solution strategies that were based on these possible explanations (PSB 4-5). Caroline's pattern of selective disclosure meant that much of the evidence for the problem, including its possible causes, was not discussed.

How does Caroline test the validity of her key beliefs?

For those five beliefs that were disclosed, Caroline checked validity by providing reasons, inquiring into Cheryl's views and reaching agreement. Once again, it is important to examine belief content to understand its relationship with validity testing. Caroline gave reasons for her views and inquired into Cheryl's reactions when she was offering support (PSB 3, 6, 8). The result was agreement on a new plan of action. But what Table 1 also shows is that Caroline did not test the relationship between her solution strategies and the possible causes of the problem (VT5) – a move that was ruled out by her decision not to disclose her view that Cheryl herself was too dependent on others' help and that the school's previous attempts to solve the problem had been inadequate (PEB 3). Nor did Caroline engage with the substantial differences that emerged between her and Cheryl.

Which values does Caroline enact in her conversation?

Caroline disclosed or withheld beliefs depending on their potential to threaten or embarrass herself, Cheryl or both. Since the principal's views had already been communicated to Cheryl at an earlier meeting, and he had authorized Caroline to follow up with Cheryl, their discussion was relatively low risk. Similarly, Caroline's offers of support (PSB 3, 6, 8) were likely to be welcomed by a teacher whom Caroline herself believed was only too ready to accept others' help (PEB 1, 2). In contrast, Caroline's own explanation of Cheryl's difficulties (PEB 1-3), which were not disclosed, had the potential to threaten the teacher and put the school and Caroline herself in a bad light.

It seems as if Caroline was in a considerable dilemma: she was concerned about the school's inept handling of Cheryl (PEB 3), while also believing that the concerns about her practice were valid (PDB 3). She was reluctant to disclose those concerns, however, because she believed that she and the principal lacked enough up-to-date observational evidence about Cheryl's classroom practice. She wrote "I already know where the principal's concerns come from but [there is] no

recorded evidence" (LHC) and later "I need to clarify and have an action [plan] to start to evidence concerns" (LHC). Could these comments suggest that Caroline's non-disclosure of her own concerns is explained by her commitment to the value of increasing the validity of the information available to both parties? Evidence gathering behaviors, however, are not, in themselves, indicative of a truth-seeking and open-minded stance. They could reflect a more persuasive and closed-minded stance if the purpose is to "evidence" undisclosed views rather than publicly test those views. While uncertainty about Caroline's enacted values remains, the combination of her non-disclosure of her critical evaluations, together with her detailed offers of support, suggest that the desire to avoid further conflict is likely the more powerful enacted value. This explanation is also suggested by the first unexpressed thought in her transcript - "I hope we don't have the same issues as last time where you [Cheryl] called in the union as you felt the principal was picking on you."

By the end of the conversation Cheryl accepted Caroline's support, (PSB 3, 6, 8) aspects of which were directly contrary to Caroline's belief that the problem was perpetuated by Cheryl's overreliance on such support (PSB 4, 5). Caroline's enacted values of avoidance of threat and embarrassment meant that much of the information of relevance to this long-standing problem was not disclosed and so little could be learned from the prior ineffective solution strategies. Caroline herself wrote at the end of her transcript, "That did not go well. I…felt it was not a 'me' conversation and I don't know why."

Liz: Private rather than Public Validity Testing

The concern selected by Liz, an assistant principal at an elementary school, involved the quality of the support given by a mentor teacher (Laura) to a provisionally registered teacher (PRT):

> My concern is with a member of my team who is the mentor of a second-year teacher. I had done the end of year evaluation with the beginning teacher and noted many areas that were of concern. We discussed some goals for the forthcoming year and discussed the areas of concern and how we were going to see growth and improvement, particularly in the area of

management. I also discussed this evaluation with the mentor teacher (Laura) and we discussed ways that she could support the beginning teacher. I have been into the beginning teacher's classroom and seen no significant shift in her practice this year. (Q, Item 1)

Liz rated her concern as 'extremely important' (Q, Item 18) because "there are children in this particular classroom who do not have an environment that is conducive to learning" (Q, Item 3). She described how she had been concerned for over a year and how she had tried to resolve the issues by meeting with Laura and the PRT, by observing the PRT's teaching and giving feedback, and through the appraisal process. She had also sought advice from the Deputy Principal and Principal. Even though she rated these prior attempts to resolve the problem as 'minimally' effective (Q, Item 16), Liz rated herself as 'very confident' in her ability to resolve it (Q, Item 19).

Which key beliefs does Liz disclose?

In her conversation with Laura, Liz disclosed both of her concerns about the PRT's performance (PDB 1-2, VT1), but not her concern about the quality of Laura's mentoring (PDB 3) (Table 2). Of her five possible explanations for the PRT's difficulties, Liz disclosed her belief about the challenges posed by a reception class (PEB 4, VT1) but not the other four hypotheses she held (PEB 1, 2, 3, 54,5, 6, 8) about her own and Laura's possible contribution to the PRT's lack of progress. She told Laura of her desire to have a three-way meeting between the PRT, Laura and herself but not of her desire to monitor the PRTS class herself. In short, Liz disclosed all her beliefs about the PRT's performance and withheld all those which were about Laura's and her own possible role in the problem.

How does Liz test the validity of her key beliefs?

When Liz did disclose her beliefs (PDB 1, 2 and PEB 4, 7) she tested them through a sequence of inquiry for Laura's reaction (VT3), gaining agreement (VT6) and then providing grounds for the agreement (VT2). For example, once Liz and Laura agreed that the PRT was making little progress

(PDB 1), the two parties consolidated their agreement by sharing observations about the PRT that confirmed their position.

[Insert Table 2 here]

Despite withholding those beliefs which could be interpreted as critical of her colleague (PDB 3, PEB 1-3), Liz still attempted to test them, for she was concerned that if the teacher was not recommended for full registration, the school would be accountable for the quality of the support it had provided. She privately rather than publicly tested her negative evaluation of the support Laura was providing by asking a series of questions about the amount and focus of Laura's mentoring and about the quality of her relationship with the PRT. She listened carefully to Laura's views, revised her own beliefs, and concluded that the support was adequate after all. She summarized by explaining to Laura that "I just wanted to have this conversation to reassure myself that we've got the right level of support in there for her, which I think, having talked to you, that we have" (RHC, lines 110-113).

One could argue that Liz's private rather than public testing was effective, for it may have saved both parties from embarrassment and threat, and Liz's initial and possibly mistaken view of the quality of Laura's mentoring was corrected. But while agreement was reached, it may not be warranted because it has been reached without critical scrutiny of the countervailing arguments about the relationship between the support provided and the PRT's lack of progress. Liz was aware of three possible explanations for the difficulty of the PRT: limitations in the mentoring provided by Laura (PEB 1-3); limitations of her own support of the mentor (PEB 5) and the special difficulties of learning to teach in a reception class (PEB 4). The decision to continue the current program was premised on the assumed validity of the third explanation and of Laura's suggestion that, with more experience, the PRT would gain confidence. While this latter explanation suggested that business as usual would be adequate, the first explanation, about the adequacy of the mentoring and support program, suggested that this solution would not be enough.

Which values does Liz enact in her conversation?

Liz's unexpressed thoughts and feelings provide clues about what motivated her selective disclosure. Of the twelve annotations she made in her LHC, seven are private attributions about how she believes Laura is experiencing the conversation. At the outset, she observes that "Laura appears quite nervous, like she is in the hot seat" (LHC, lines 11-13). Half way through, Liz is thinking that "Laura is starting to relax a bit more now" (LHC, lines 56-57) and at the end of the conversation Liz observes that she "feels I've got her onside" (LHC, lines 117-118) and that they have reached a "consensus" (LHC, line 122). Her careful monitoring of Laura's emotional reactions suggests that Liz was concerned to avoid negative affect and thus was not prepared to disclose those beliefs that might cause upset. The result was that Liz and Laura agreed on the nature of the problem (PDB 1-3) and on how to resolve it (PSB 1), without discussing the explanation that Liz had written about in her questionnaire – the inadequacy of the support provided by Laura herself (PDB 3).

As for Caroline, Liz's desire to avoid threat or embarrassment to either party seems to explain what was and was not disclosed and publicly tested. Criticisms of the performance of the PRT were shared, while beliefs about Laura's performance and self-criticism were not. The explanation of this pattern is not that Liz does not want to communicate critical evaluations, for she does in some instances (PDB 1-2, VT 1-2). Rather, it is that the leader does not want or is unable to express criticisms that are about the parties in the conversation (PDB 3, PEB 1-3).

Paul: Open Behaviors Masking Closed-mindedness

Paul, one of two deputy principals of a large high school, was concerned about the strained relationship between himself and his fellow deputy Gary. Paul was responsible for a pastoral care team made up of the heads of each grade level called Heads of Year (HoY). Gary was in charge of the Heads of Faculty (HoF), each of whom was responsible for a particular learning area. Paul wrote about his concern in the questionnaire as follows:

After a recent Heads of Faculties meeting, which Gary chairs, one of the Heads of Faculty who attended the meeting approached me afterwards saying he was concerned at several comments made by Gary. I was not at the meeting, but was told that Gary allegedly said: "The Heads of Faculty (HoF) are visible around the school but the Heads of Year (HoY) are not. They are too often in their office. The HoYs will not be helping with teacher relief but the senior leadership team will." My concern is - what was Gary's motive in making these comments in that setting? Was Gary deliberately trying to be destructive and cause divide amongst the two core groups of middle school management? (Q, Item 1)

Paul saw Gary as continuing a pattern of divisive and competitive behavior which he believed other school leaders had noticed but had failed to address. He had tried to talk to Gary on four occasions and rated these attempts to resolve his concern as 'very difficult' (Q, Item 15) and only 'minimally effective' (Q, Item 16). Paul believed it was 'extremely important' to solve the problem (Q, Item 18), but he was only 'somewhat' confident that he could do so (Q, Item 19).

Which key beliefs does Paul disclose?

Paul disclosed his three concerns (PDB 1-3, VT1) and the grounds for them (VT2), in a way that matched how he had described them in his questionnaire. Unlike Caroline and Liz, he was prepared to give negative feedback to his colleague and discuss his belief about Gary's motives for his public criticism (PEB 1). He withheld his more general attributions about Gary's character (PEB 2, 4) and his own prior failure to provide clear feedback about Gary's behavior (PEB 3).

[Insert Table 3 here]

How does Paul test the validity of his key beliefs?

Paul disclosed his concerns (PDB 1-3, VT1) along with the grounds for those concerns (VT2), without assuming their truth. His qualifier "what you reportedly said" (RHC, line 3) acknowledged the possibility that his leadership of the meeting had been misreported and created space for Gary to disagree. His perspectival language is also seen in the way he communicated his belief that Gary

was motivated by a desire to cause division ("I feel", RHC line 7; "It looks like it is divisive", RHC, line 94). An open-minded stance is further suggested by his direct inquiry for Gary's reaction (PEB 1, VT3) to his attribution "Now, what do you think about what I said around the divide and whether that's intentional?" (RHC, lines 104-5).

These behaviors were not sufficient, however, to gain agreement about the nature and cause of the problem. Gary expressed differing views about all of them and the differences remained unresolved. Paul decided not to engage with Gary's criticism of his team (PDB 3), choosing instead to probe and listen and then, when he had had enough, find a way to stop the conversation and the public criticism. Despite uncertainty about whether the public criticism occurred, and disagreement about its justification, Paul and Gary did agree on a strategy to prevent any future public airing of such criticism (PSB 1, VT6).

While the validity testing pattern for Caroline and Liz can be described as *disclose what is safe, provide grounds, and gain agreement*, the pattern for Paul can be summarized as *disclose what may be unsafe, provide grounds, inquire and gain disagreement*. What these three leaders have in common is difficulty using disagreement to publicly test the validity of competing views before agreeing on a course of action (VT4). What is arguably the most powerful of all the validity testing behaviors was not used by any of the three leaders.

Which values does Paul enact in his conversation?

Caroline and Liz were motivated by a desire to avoid negative emotion and thus withheld beliefs that were self-critical and critical of the other party. Paul, by contrast, seemed to be motivated more by the desire to finally confront his colleague and get what he wanted, than by avoiding upset.

While Paul's perspectival speech and direct inquiry for Gary's reactions are suggestive of an open-minded stance, his undisclosed thoughts cast considerable doubt on the sincerity of these behaviors. For example, when Gary denied being divisive and gave an alternative explanation for

27

his behavior, Paul judged him privately as "either a liar or very naive" (LHC, line 105). It may be that Paul had good reason to doubt the veracity of Gary's statements. He described in his questionnaire that "I also have significant doubt of [sic] Gary's overall integrity in telling and reporting the truth" (Q, Item 2). Public testing of one's own beliefs does not entail automatic acceptance of others' denials and alternative explanations. Instead, what the search for truth requires is respectful disclosure and checking of the reasons for skepticism. Paul is unlikely to do this while he experiences glee at finally confronting Gary (LHC, line 15) and delights at catching him out (LHC, line 17). Paul's unexpressed thoughts and feelings suggest that his primary motivation was to win by shutting down Gary's criticism rather than to understand and resolve Gary's concerns about the work of his team.

Discussion

The proposed model of quality problem-solving developed in this study contributes to an integration of the rational and more intuitive theories of problem-solving. By anchoring the quality of problem-solving conversations in the values and behaviors associated with public validity testing, we address the criticism levelled at rational theories that they assume a linear sequence of problem identification, analysis and solving that is mismatched to the fast pace, complexity and uncertainty of leadership work. Our process model does not require that everything is tested – only key beliefs- and nor does it require that they be tested in a particular sequence. The alignment of problem cause with solution strategies, for example, may be tested either before or after solutions are proposed. What is critical, for the ethical and educational reasons discussed earlier, is that they are tested before commitment of substantial financial and human resources. Our model also addresses the error prone nature of intuitive problem solving by specifying the values and behaviors involved in publicly checking, at any stage of the problem-solving conversation, the validity of key beliefs about a problem.

While generalizations to other experienced educational leaders are not warranted, our findings suggest, that the stance of our three leaders was more that of truth claiming than truth seeking, for they did not disclose many of their most important beliefs and used a limited range of validity testing behaviors. This finding closely parallels that of Mintrop and Zumpe (2019) who used a very different methodology from our own, but crucially, studied the problem-solving of experienced leaders in their real on the job contexts.

Our classification of belief content showed that patterns of disclosure and non-disclosure were systematic, with leaders' beliefs about problem explanations (PEB) far less likely to be disclosed (2 of 12), than their beliefs about the nature of a problem (PDB; 7 of 12) or how it could be solved (PSB; 5 of 11). In short, there was a considerable difference between what our leaders believed about their problem and what they were prepared to talk about. The explanation for this pattern of disclosure is not a paucity of causal reasoning on the part of the leaders, for our probes into their unexpressed thoughts revealed rich, but not necessarily accurate, beliefs about the causes of the problem. A more likely explanation is that they believed their disclosure would have risked evoking the negative emotions which their questionnaires and LHC indicated they wished to avoid (Yariv, 2006). Since discussion of solution strategies posed less risk of negative emotion, the desire to avoid such emotion is a likely explanation of the predominance of solution talk over causal talk.

A second possible explanation for the predominance of solution talk over causal talk lies in leaders' beliefs about what counts as a quality conversation. Some of the unexpressed thoughts of each leader suggest that they saw moving swiftly to a solution as a sign of a good conversation. For Paul, it was about getting quickly to what he wanted: "I'm going to try and speed this up and just put what I want on the table" (LHC, lines 123-125). For both Caroline and Liz, a quality conversation was one in which they 'looked forward'. Caroline wrote, "Why did I ask that, as I know what she will bring up again, I've reopened the conversation to repeat and not look forward" (LHC, lines 121-123). Similarly, Liz annotated her transcript with, "I've done a lot of listening and

nodding, [and] need to put my opinion forward to move things along" (LHC, lines 78-82). This evidence, like that found by an earlier study of principal problem-solving (Brenninkmeyer, & Spillane, 2008), suggests that leaders' beliefs about what counts as a quality conversation may be contributing to their 'quick fix' pattern of problem-solving.

When leaders did disclose their beliefs, they employed a limited range of validity testing behaviors, with provision of grounds for their beliefs, inquiry into the other's views and gaining of agreement, the predominant behaviors. Even though each leader encountered difference and disagreement, none was able to harness the powerful opportunity that disagreement provides for examining the validity of differing views. Instead, leaders' thoughts indicated that difference was treated as an obstacle to acceptance of one's own views – a stance that is contrary to the open-minded disposition required for public validity testing (Hare, 2003). Also absent was examination of the alignment between problem explanations and proposed solutions.

Although the interpersonal values in our model of validity testing are widely espoused (Argyris & Schön, 1996), there was little evidence of their enactment in these conversations. For Liz and Caroline, the stance was a genuine desire to gain agreement to a course of action that was based on their private preconception of what was needed. They succeeded in avoiding conflict and maintaining relationships but were unable to integrate these values with what is required for public validity testing. The result was solutions that were misaligned to privately held beliefs about what was really required and continuation of the same solution strategies that had thus far been unsuccessful. A limitation of our study is that we could not follow-up these conversations to establish whether the agreed solutions were implemented and effective.

The conclusion we draw from leaders' questionnaires, unexpressed thoughts and feelings (LHC) and from their actual speech, is that leaders' problem-solving processes were motivated more by relational than epistemic values. When key beliefs were seen as potentially threatening or embarrassing to self or others, they were withheld, unless the relationship had already broken down,

as was arguably the case for Paul and Gary. The other two leaders tended not to disclose and publicly test those beliefs that they anticipated would cause embarrassment or threat to themselves or others, regardless of the centrality of those beliefs to their understanding and resolution of problem. When leaders frame their beliefs in ways that are judgmental, absolutist and critical, they experience a dilemma between disclosing their views and protecting the relationship (Argyris, 1999). Rather than avoiding the dilemma by reframing their beliefs so they could be respectfully tested, leaders focused instead on safer beliefs, testing their 'risky' beliefs either in private or not at all. The dilemma between being respectful and open that was felt by two of the three leaders is not inevitable. If leaders frame their beliefs as specific observations rather than as generalized attributions, and as explanatory hypotheses rather than as certainties, then it would be easier for them to publicly test their validity. Such framing, consistent with the three values in the model, recognizes the fallibility of beliefs, the complex and contested nature of organizational problems, and is respectful in the sense of being open to difference (Hare, 2003).

Implications for Research on Problem-Solving Conversations

This study has contributed a process model of quality problem-solving that integrates the content of beliefs about a problem with the epistemic and relational values and behaviors involved in the public testing of those beliefs. Our focus on problems sourced in the on-the-job responsibilities of our leaders, provides an authentic context in which the influence of problem history, work relationships and real world consequences can be studied. We have provided rich qualitative data showing the strong interaction between the relational and task aspects of problem-solving conversations and how the former trumps the latter in ways that sacrifice the quality of the problem-solving. Our study also contributes through the development of a methodology which enables belief content to be linked to particular validity testing behaviors, and this has revealed the differential disclosure and validity testing of different types of belief. In addition, our probing of the private thoughts of our leaders has advanced our understanding of why certain beliefs are not disclosed and tested.

31

The inclusion of data in this study about leaders' undisclosed thoughts as well as their behavior has shown that each source provides differing insights into leaders' validity testing. At times, leaders will use behaviors associated with public testing, while privately maintaining a closed and defensive stance. As Argyris and Schön argued (1974), it is the disposition to inquire and learn that is central to the pursuit of validity. Behavioral measures on their own, as with Paul, may not give a reliable indication of such a disposition. Both types of evidence are needed. Behavioral measures are essential if one is to identify the practices that may be involved in quality problem-solving conversations and thus build leadership capability. Cognitive measures help researchers and leaders access, evaluate and improve the forms of reasoning that drive these behaviors.

Further research is needed with larger samples drawn from different leadership domains to test the extent to which the patterns of validity testing used by our educational leaders are typical of experienced leaders in other types of organization. Quantitative research, using path analytic techniques which test the full model depicted in Figure 1, would further our understanding of the interaction between the relational and epistemic components of quality problem-solving. Follow-up studies of the outcomes of the action taken subsequent to the conversations are also needed to test the validity of the model.

Implications for Leadership Practice and Development

The ways in which leaders resolve team and organizational problems has powerful consequences for themselves, their organizations and the communities they serve. This study has shown how experienced leaders may assume rather than publicly test the validity of their beliefs about the nature, causes and resolution of the problems they are responsible for resolving, especially when they experience a tension between such testing and maintaining their relationships. Given the high percentage of problems that pose such emotional risk, it is important for leaders and leadership developers to understand how this risk may be mitigated in ways that foster more quality problem-solving (Le Fevre, Sinnema, & Meyer, 2016; Robinson, 2016).

This study, together with a few already published intervention studies, suggests the types of professional learning experience that may help leaders to navigate this tension (Avolio et al., 2009; Marcy & Mumford, 2010; Riggio & Lee, 2007). First, the learning should be situated in authentic emotionally-laden conversations in which leaders are likely to experience a tension between being respectful and being open about their beliefs. Second, opportunities to evaluate and revise how beliefs are formulated are important, as disrespectful thoughts prevent respectful disclosure of key beliefs about the problem. Third, the learning opportunity should be informed by evidence, such as annotated transcripts, which makes leaders' thoughts and speech visible and provides insight into their interpersonal values. Fourth, systematic follow-up of the professional learning should include evaluation of problem outcomes on the trained problems and generalization of leaders' learning to untrained problems.

Rather than a course on problem-solving, or a course on relationships, the findings of this study suggest leaders need help with how to integrate relationship building with problem-solving in the context of their own emotionally laden on-the-job problems. The challenge for leaders is not to eliminate their preconceptions, for they are an inevitable feature of their memory and information processing capabilities, but to eliminate the assumption that their preconceptions are correct.

References

Argyris, C. (1976). Single-loop and double-loop models in research on decision making. *Administrative Science Quarterly*, 21(3), 363-375. doi: 10.2307/2391848

Argyris, C. (1999). *On organizational learning* (2nd ed.). Oxford, UK: Blackwell.

Argyris, C. (1990). *Overcoming organizational defenses: Facilitating organizational learning.* Boston, MA, Allyn and Bacon.

Argyris, C. (2010). *Organizational traps: Leadership, culture, organizational design*. Oxford: Oxford University Press.

Argyris, C., & Schön, D. (1974). *Theory in practice: Increasing professional effectiveness.* San Francisco, CA: Jossey-Bass.

Argyris, C., & Schön, D. (1996). *Organizational learning II: Theory, method and practice.* Reading, MA: Addison Wesley.

Allison, D. J., & Allison, P. A. (2003). Both ends of a telescope: Experience and expertise in principal problem solving. *Educational Administration Quarterly*, 29(3), 302-322.

Avolio, B. J., Reichard, R. J., Hannah, S. T., Walumbwa, F. O., & Chan, A. (2009). A meta-analytic review of leadership impact research: Experimental and quasi-experimental studies. *The Leadership Quarterly, 20*(5), 764-784. doi: org/10.1016/j.leaqua.2009.06.006

Ball, D. (2018). *The dreams, possibilities, and necessity of public education*. Paper presented at the Presidential Address for the Annual Meeting of the American Educational Research Association (AERA), in New York, NY (13-17 April).

Bambacas, M., & Patrickson, M. (2008). Interpersonal communication skills that enhance organisational commitment. *Journal of Communication Management, 12*(1), 51-72. doi: 10.1108/13632540810854235

Bokeno, R. M. (2003). The work of Chris Argyris as critical organization practice. *Journal of Organizational Change Management*, *16*(6), 633-649.

Bowen, G. A. (2008). Naturalistic inquiry and the saturation concept: A research note. *Qualitative Research*, 8(1), 137-152. doi: 10.1177/1468794107085301

Bransford, J. D., & Stein, B. S. (1993). *The IDEAL problem solver*. Centers for Teaching and Technology.

Brenninkmeyer, L. D., & Spillane, J. P. (2008). Problem-solving processes of expert and typical school principals: A quantitative look. *School Leadership and Management, 28*(5), 435-468. doi: 10.1080/13632430802517209

Bryk, A. S., Gomez, L. M., Grunow, A., & Le Mahieu, P. G. (2015). *Learning to improve: How America's schools can get better at getting better.* Cambridge, MA: Harvard Education Press.

Colquitt, J. A., Scott, B. A., & LePine, J. A. (2007). Trust, trustworthiness, and trust propensity: A meta-analytic test of their unique relationships with risk taking and job performance. *Journal of Applied Psychology, 92*(4), 909–927. doi: 10.1037/0021-9010.92.4.909

DeCuir-Gunby, J. T., Marshall, P. L., & McCulloch, A. W. (2011). Developing and using a codebook for the analysis of interview data: An example from a professional development research project. *Field methods*, 23(2), 136-155.

Dempster, N., & Berry, V. (2003). Blindfolded in a minefield: Principals' ethical decision-making. *Cambridge Journal of Education*, *33*(3), 457-477.

Dewulf, A., & Bouwen, R. (2012). Issue framing in conversations for change: Discursive interaction strategies for "doing differences". *The Journal of Applied Behavioral Science, 48*(2), 168-193.

Fishbein, M., & Azjen, I. (1975). *Belief, intention and behavior: An introduction to theory and research.* Reading, MA: Addison Wesley.

Hare, W. (2003). The ideal of open-mindedness and its place in education. *Journal of Thought, 38*(2), 3-10. doi: jstor.org/stable/42589737

Hart, W., Albarracin, D., Eagly, A., Brechan, I., Lindberg, M. J., & Merrill, L. (2009). Feeling

validated versus being correct: A meta-analysis of selective exposure of information.

Psychological Bulletin, 135(4), 555 - 588. doi: 10.1037/a0015701

Kahneman, D. (2011). *Thinking, fast and slow.* New York, NY: Farrar, Straus and Giroux.

Le Fevre, D., Sinnema, C., & Meyer, F. (2016). Possibilities for school improvement. In V. M. J.

Robinson, D. M. Le Fevre, & C. Sinnema (Eds.), *Open-to-learningTM leadership: How to build*

trust while tackling tough issues (pp. 231– 245). Melbourne, VIC: Hawker Brownlow Education.

Le Fevre, D. M., & Robinson, V. M. J. (2015). The interpersonal challenges of instructional

leadership: Principals' effectiveness in conversations about performance issues. *Educational*

Administration Quarterly. doi:10.1177/0013161x13518218

Leithwood, K., & Stager, M. (1989). Expertise in principals' problem solving. *Educational*

Administration Quarterly, 25(2), 126-161. doi: 10.1177/0013161X89025002003

Leithwood, K., & Steinbach, R. (1995). *Expert problem solving: Evidence from school and district*

leaders. Albany, NY: State University of New York Press.

Leithwood, K. A., Steinbach, R., & Raun, T. (1993). Superintendents' group problem-solving

processes. *Educational Administration Quarterly*, 29(3), 364-391.

MacQueen, K., McLellan-Lemal, E., Bartholow, K., & Milstein, B. (2008). Teambased codebook

development: Structure, process, and agreement. In G. Guest & K. M. MacQueen (eds.)

Handbook for team-based qualitative research (pp. 119–35). Lanham, MD: AltaMira.

Marcy, R. T., & Mumford, M. D. (2010). Leader cognition: Improving leader performance through

causal analysis. *The Leadership Quarterly, 21*(1), 1-19. doi: org/10.1016/j.leaqua.2009.10.001

Mazutis, D., & Slawinski, N. (2008). Leading organizational learning through authentic dialogue.

Management Learning, 39(4), 437.

Milkman, K. L., Chugh, D., & Bazerman, M. H. (2009). How can decision making be improved?

Perspectives on Psychological Science, 4(4), 379-383. doi:10.1111/j.1745-6924.2009.01142.x

Mintrop, R., & Zumpe, E. (2019). Solving real-life problems of practice and education leaders' school improvement mind-set. *American Journal of Education*, *125*(3), 295-344.

Meyer, F., Sinnema, C., & Patuawa, J. (2019). Novice principals setting goals for school improvement in New Zealand. *School Leadership and Management*, *39*(2), 198–221. doi:10.1080/13632434.2018.1473358

Morse, J. (2015). Critical analysis of strategies for determining rigor in qualitative inquiry. *Qualitative Health Research*, 25(9), 1212-1222. doi: 10.1177/1049732315588501

Morse, J., & Field, P. (1995). *Qualitative methods for health professionals* (2nd edition ed.). Thousand Oaks, CA: Sage.

Mumford, M. D., Gibson, C., Giorgini, V., & Mecca, J. (2014). Leading for creativity: People, products and systems. In D. V. Day (Ed.). *The Oxford handbook of leadership and organizations"* (pp. 757-782). Oxford, UK: Oxford University Press.

Mumford, M. D., Todd, E. M., Higgs, C., & McIntosh, T. (2017). Cognitive skills and leadership performance: The nine critical skills. *The Leadership Quarterly, 28*(1), 24-39. doi: 10.1016/j.leaqua.2016.10.012

Mumford, M. D., Watts, L. L., & Partlow, P. J. (2015). Leader cognition: Approaches and findings. *The Leadership Quarterly, 26*(3), 301-306.

Mumford, M. D., Zaccaro, S. J., Harding, F. D., Jacobs, T. O., & Fleishman, E. A. (2000). Leadership skills for a changing world: Solving complex social problems. *The Leadership Quarterly, 11*(1), 11-35. doi: 10.1016/S1048-9843(99)00041-7

Nickles, T. (1981). What is a problem that we might solve it? *Synthese, 47*(1), 85-118.

O'Reilly, M., & Parker, N. (2012). 'Unsatisfactory Saturation': A critical exploration of the notion of saturated sample sizes in qualitative research. *Qualitative Research*, 13(2), 190-197. 10.1177/1468794112446106

Phillips, D. C. (1987). Validity in qualitative research: Why the worry about warrant will not wane. *Education and Urban Society 20*(1), 9-24. doi: 10.1177/0013124587020001003

Popper, K. (1999). *All life is problem solving.* London, UK: Routledge.

Richardson, V. (1996). The role of attitudes and beliefs in learning to teach. In J. Sikula (Ed.).

Handbook of research on teacher education (2nd ed.,pp. 102-119). New York: MacMillan.

Riggio, R. E., & Lee, J. (2007). Emotional and interpersonal competencies and leader development.

Human Resource Management Review, 17(4), 418-426. doi: org/10.1016/j.hrmr.2007.08.008

Robinson, V. M. J. (2014). Single and double loop learning. In D. Phillips (Ed.), *Encyclopaedia of*

Educational Theory and Philosophy (pp. 755–757). Thousand Oaks, CA: Sage.

Robinson, V. M. J. (2016). The theory and practice of open-to-learningTM Leadership. In V. M. J.

Robinson, D. M. Le Fevre, & C. Sinnema (Eds.), *Open-to-learningTM leadership: How to build*

trust while tackling tough issues (pp. 231–245). Melbourne, VIC: Hawker Brownlow Education.

Robinson, V. M. J., Sinnema, C. E., & Le Fevre, D. (2014). From persuasion to learning: An

intervention to improve leaders' response to disagreement. *Leadership and Policy in Schools*,

13(3), 260–296. doi:10.1080/15700763.2014.922997

Saiti, A. (2015). Conflicts in schools, conflict management styles and the role of the school leader:

A study of Greek primary school educators. *Educational Management Administration &*

Leadership, *43*(4), 582-609.

Schön, D. A. (1991). *Educating the reflective practitioner*. San Francisco, CA: Jossey Bass.

Simon, H. A. (1993). Decision-making: Rational, non-rational and irrational. *Educational*

Administration Quarterly, 29(3), 392-411.

Sinnema, C. E. L., Le Fevre, D., Robinson, V. M. J., & Pope, D. (2013). When others' performance

just isn't good enough: Educational leaders' framing of concerns in private and public.

Leadership and Policy in Schools, 12(4). doi:10.1080/15700763.2013.857419

Sinnema, C. E. L., & Robinson, V. M. J. (2012). Goal setting in principal evaluation: Goal quality

and predictors of achievement. *Leadership and Policy in Schools*, 11(2), 135–167.

doi:10.1080/15700763.201

Spiegel, J. S. (2012). Open-mindedness and intellectual humility. *Theory and Research in Education, 10*(1), 27-38. doi:10.1177/1477878512437472

Spillane, J. P., White, K. W., & Stephan, J. L. (2009). School principal expertise: Putting expert-aspiring principal differences in problem solving processes to thet test. *Leadership and Policy in Schools*, 8(2), 128-151. doi: 10.1080/15700760902737188

Stanovich, K. E., & West, R. F. (1997). Reasoning independently of prior belief and individual differences in actively open-minded thinking. *Journal of Educational Psychology, 89*(2), 342-357. doi: 10.1037/0022-0663.89.2.342

Sternberg, R. J., & Horvarth, J. A. (1995). A prototype view of expert teaching. *Educational Researcher*, 24(6), 9-17.

Sun, J., & Leithwood, K. (2015). Direction-setting school leadership practices: A meta-analytical review of evidence about their influence. *School Effectiveness and School Improvement*, *26*(4), 499-523.

Tjosvold, D. (2008). The conflict-positive organization: It depends upon us. *Journal of Organizational Behavior, 29*(1), 19-28. doi: 10.1002/job.473

Tjosvold, D., Sun, H. F., & Wan, P. (2005). Effects of openness, problem solving, and blaming on learning: An experiment in China. *The Journal of Social Psychology*, 145(6), 629-644. doi: 10.3200/SOCP.145.6.629-644

Yariv, E. (2006). Mum effect: Principals' reluctance to submit negative feedback. *Journal of Managerial Psychology*, *21*(6), 533 – 546. doi: 10.1108/02683940610684382

Zaccaro, S. J., Mumford, M. D., Connelly, M. S., Marks, M. A., & Gilbert, J. A. (2000). Assessment of leader problem-solving capabilities. *The Leadership Quarterly, 11*(1), 37-64. doi:10.1016/S1048-9843(99)00042-9

Zand, D. E. (2016). Reflections on trust and trust research: Then and now. *Journal of Trust Research*, (6)1, 63-73, doi: 10.1080/21515581.2015.1134332
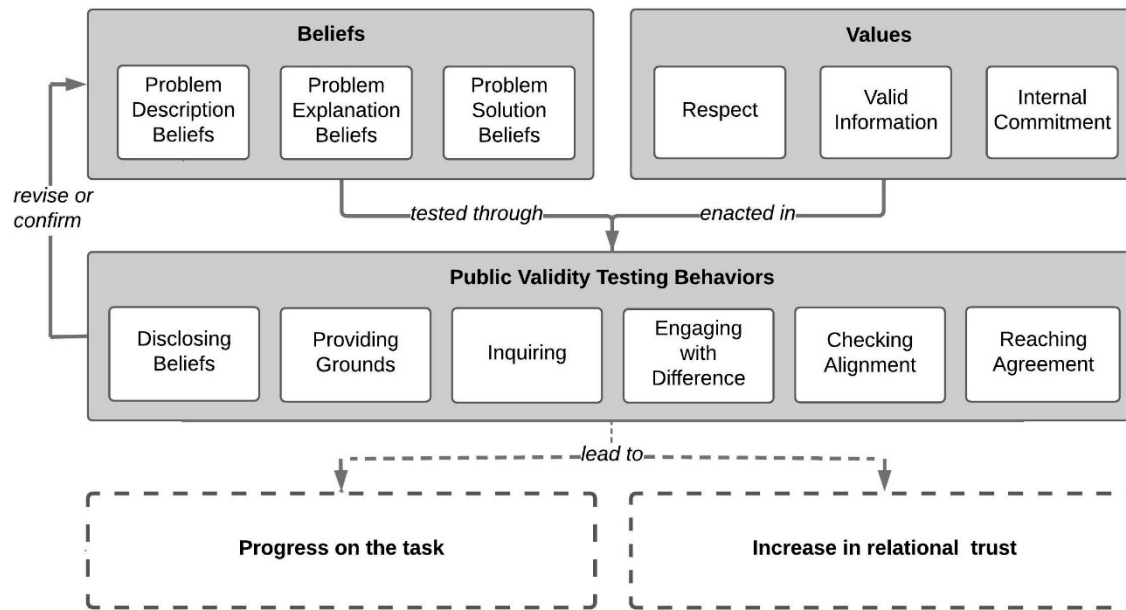
Figure 1: *A process model of quality problem-solving*

Table 1.

Caroline's Key Beliefs and Validity Testing Behaviors in Conversation with Cheryl (Teacher)

Key beliefs	Presence of validity testing behaviors*					
Caroline believes that…	VT1	VT2	VT3	VT4	VT5	VT6
Problem description beliefs (PDB)						
1. the principal is concerned about Cheryl's loud voice (Q, RHC)	●		●			
2. the principal is concerned about the impact of Cheryl's loud voice on classroom noise (Q, RHC)	●		●			
3. some of the principal's concerns are valid (Q, RHC)						
4. evaluators from the school review agency and math facilitator are concerned about Cheryl's loud voice (Q, LHC)						
5. same type of issue was raised two years earlier (Q, LHC)						
6. the issue might be hard to address as recent observational evidence is not available (Q)						
Problem explanation beliefs (PEB)						
1. Cheryl is too dependent on others' help (Q)						
2. Cheryl relied too much on her support in the past (Q)						
3. the support program for Cheryl is not sufficiently robust or consistent (Q)						
Problem solution beliefs (PSB)						
1. she needs to put a monitoring system in place to gain evidence about extent of problem (Q)						
2. after monitoring, she needs to give feedback to Cheryl (Q)						●
3. she can be a support person (RHC)	●	●	●			●
4. she should be more than a support person as she wants to make changes (LHC)						
5. Cheryl should take more ownership of her problems (LHC)						
6. she will talk with the principal about the basis of his concerns (RHC)	●	●	●			●
7. she knows basis of principal's concerns (LHC)						
8. she will liaise with the teachers for whom Cheryl is relieving to clarify work expectations (RHC)	●	●	●			●

*VT1 = Disclosing key beliefs; VT2 = Providing grounds; VT3 = Inquiring; VT4 = Engaging with difference; VT5 = Checking alignment; VT6 = Reaching agreement

Table 2.

Liz's Key Beliefs and Validity Testing Behaviors in Conversation with Laura (Tutor Teacher)

Key beliefs	Presence of validity testing behaviors*					
Liz believes that….	VT1	VT2	VT3	VT4	VT5	VT6
Problem description beliefs (PDB)						
1. the provisionally registered teacher (PRT) is showing little progress in her classroom management (Q, RHC)	●	●	●			●
2. PRT does not communicate with colleagues in professional conversations or meetings (RHC)	●	●	●			●
3. the support provided by Laura for PRT is not adequate (Q)			●			●
Problem explanation beliefs (PEB)						
1. Laura is an inexperienced tutor teacher and is finding the role difficult (Q)						
2. Laura is too gentle and not proactive enough (Q)			●			
3. Laura and PRT have a difficult working relationship (Q)			●			
4. the reception class is a difficult context for any PRT (RHC)	●	●				●
5. she has not had sufficient meetings with Laura and the PRT together (Q)						
Problem solution beliefs (PSB)						
1. a three-way meeting is needed to check PRT's understanding of her recent evaluation (RHC)		●	●			●
2. she needs to closely monitor the progress of the PRT's class (Q)						

*VT1 = Disclosing key beliefs; VT2 = Providing grounds; VT3 = Inquiring; VT4 = Engaging with difference; VT5 = Checking alignment; VT6 = Reaching agreement

Table 3.

Paul's Key Beliefs and Validity Testing Behaviors in Conversation with Gary (Deputy Principal)

Key beliefs	Presence of validity testing behaviors*					
Paul believes that…	VT1	VT2	VT3	VT4	VT5	VT6
Problem description beliefs (PDB)						
1. Gary is reported to have publicly criticized his team (Q, RHC)	●	●				
2. public criticism is destructive (Q, RHC)	●	●				
3. Gary has issues with his team (Q, RHC)	●		●			
Problem explanation beliefs (PEB)						
1. Gary is deliberately trying to cause harm and be divisive (Q, LHC, RHC)	●	●	●			
2. Gary avoids direct conflict so manipulates to get what he wants (Q)						
3. he has not been clear enough about his concerns about Gary's behavior (Q)						
4. Gary disregards importance of team work (Q)						
Problem solution beliefs (PSB)						
1. Gary should give him private 'heads up' about concerns and only express them in vague terms (RHC)	●		●			●

*VT1 = Disclosing key beliefs; VT2 = Providing grounds; VT3 = Inquiring; VT4 = Engaging with difference; VT5 = Checking alignment; VT6 = Reaching agreement