The Influence of Collaboration on Program Outcomes: The Colorado NurseFamily Partnership

Darrin Hicks, Carl Larson, Christopher Nelson, David L. Olds and Erik Johnston

Eval Rev 2008; 32; 453 originally published online Mar 17, 2008; DOI: 10.1177/0193841X08315131

The online version of this article can be found at:
http://erx.sagepub.com/cgi/content/abstract/32/5/453
The Influence of Collaboration on Program Outcomes
The Colorado Nurse–Family Partnership

Darrin Hicks
Carl Larson
Christopher Nelson
University of Denver
David L. Olds
University of Colorado
Erik Johnston
Arizona State University

Though collaboration is often required in community initiatives, little evidence documents relationships between collaboration and program success. The authors contend that clarification of the construct collaboration is necessary for investigating its contribution to the success of community initiatives. After respecifying collaboration, they present a study of a multisite program that involved varying degrees of collaboration in the 16 communities adopting a nurse home visitation program. The authors employ hierarchical linear modeling (HLM) to test the predictive power of individual participant characteristics and examine the increased accuracy of predictions from a second level model of site qualities—specifically, features of the collaborative process associated with different sites. The first-level model predicted approximately 10% of the variance in attrition, or dropout, of program clients. The second-level model accounted for an additional 28% of the variance in attrition. A theory of commitment transfer is offered as a first explanation of this result.

Keywords: collaboration; process quality; community health; nurse–family partnership
Collaborative partnerships—alliances among community stakeholders and organizations from multiple sectors working together to improve conditions that promote and sustain community health—are an increasingly prominent strategy for addressing community health needs (Feinberg et al. 2004; Berkowitz 2001; Roussos and Fawcett 2000). Federal and state health agencies, Healthy People 2010, the Institute of Medicine, and the National Public Health Performance Standards all have explicitly supported the use of collaborative partnerships (Zahner 2005). Private philanthropic trusts are building in community collaboration as a design feature of the projects that they fund (Backer 2003). The intrinsic value of collaboration in improving community health has become taken for granted by the public health community.

Despite the extensive resources invested in collaborative efforts, there is little empirical support demonstrating their effectiveness in improving community health. Most evaluations of collaborative partnerships are partial and non-systematic with few of them attempting to measure the relationship between collaboration and health outcomes (Backer 2003; Berkowitz 2001). This is not to say that extant evaluations of community collaboration are insignificant. Berkowitz’s (2001) review of community-based coalitions cites many careful case studies that highlight the substantial value of community collaboration. However, these case studies often focus solely on the constitutive conditions of effective coalition functioning. The study of community coalitions, hence, is largely a “wisdom literature” that, although offering much of great value to those who design and maintain community coalitions, rarely directly examines the relationship between collaboration and program outcomes.

The results of those evaluations that do measure the health outcomes of collaborative partnerships provide only marginal evidence of the effectiveness of collaborative efforts (Roussos and Fawcett 2000; Kreuter et al. 2000). This marginal evidence of effectiveness may be because, as Kreuter et al. (2000) have suggested, we may not be asking the right questions when we evaluate collaborative efforts. That is, improvements in health status and changes in health systems may occur as the result of collaboration but go undetected because neither current conceptualizations of collaboration nor the instruments designed to measure its effects are sufficient to establish the relationship between the quality of collaboration and health outcomes.

**Authors’ Note:** The authors wish to thank Dr. Ruth O’Brien from the National Center for Children and Families and Jennifer Atler from Invest in Kids for their assistance. The Colorado Trust supported this research.
We concur with Kreuter et al.’s (2000) claim that more sophisticated accounts of the nature of collaboration and instruments for evaluating the effects of collaboration are necessary: a claim that has also been forwarded by Berkowitz (2001) and Yin and Kaftarian (1997). Despite the conflicting results of contemporary research on the success of collaboration, we assume that collaboration will continue to be used as a mechanism for improving community health: As Richardson and Allegrante (2000) claimed, “We need partnerships because most of the problems we will face in the 21st century will require multisectoral, multidisciplinary, and multicomponent efforts” (p. 375). If collaboration is to contribute to the success of community health initiatives, we must increase the precision in evaluating its properties and effects.

**Reconceptualizing Collaboration**

**Collaboration as Communication**

The first step toward developing more nuanced means to measure the relationship between collaboration and program outcomes is refining the definition of the concept. This begins by distinguishing collaboration from coordination. Chrislip and Larson (1994) argued that collaboration entails more than sharing information and transferring knowledge, and more important, it entails more than coordinating efforts so each party can achieve its goals. Rather, the aim of collaboration “is to create a shared vision and joint strategies to address concerns that go beyond the purview of any particular party” (Chrislip and Larson 1994, 5).

Chrislip and Larson’s (1994) account of collaboration is consistent with Butterfoss et al.’s (1993) argument that the agreement to work together and the coordination of services toward a shared goal are necessary but not sufficient features of coalitions. The “defining characteristic” of coalitions, in comparison to other types of groups, is the shift from advocating for parochial interests to understanding the problem and advocating for solutions from the vantage of the coalition itself (Butterfoss et al. 1993, 316). Collaboration, then, is more than the coordination of effort toward mutually beneficial ends. Coordination involves directing the effort of multiple agents in the most efficient and effective manner to accomplish mutually known ends. Following Heifetz (1994), we would contend that when there is a clear consensus on the definition of the problem and how to solve it, coordination, not collaboration, might be the best strategy for proceeding. Similarly, if stakeholders must choose between two relatively clear options, they would be better served by a third-party process that can help them decide (mediation).
or decide for them (arbitration or litigation). Collaboration, on the other hand, is a communicative activity in which stakeholders invent new ways of seeing and responding to social problems. Collaboration begins without full knowledge of the nature of the problem or solution. Stakeholders must engage each other in a serious process of mutual understanding and problem solving to create possible solutions. Coordination begins once stakeholders have clearly defined the problem, set target goals, established a plan for reaching those goals, and reached agreement as to the roles each stakeholder will take in executing the plan. If there is not a clear understanding of the problem, solution, roles, or means of executing the plan to reach targeted goals, stakeholders should engage in collaboration prior to coordinating their efforts. Coordination, hence, is a necessary but not sufficient feature of truly “collaborative partnerships.”

It is not wholly surprising that the presence of multi-agent coordination often does not lead to demonstrable improvements, at least not any more than efforts of single agents. In an examination of six large-scale community health initiatives that built in “collaboration” as a design feature of the initiative, ranging from palliative care to teen pregnancy prevention, Larson and Hicks (2004) found clear differences in the ability to impact health status and health systems in those initiatives for which stakeholders made clear delineations between collaborative and coordinative activities. The three initiatives for which stakeholders collaboratively established a shared community vision and joint strategies for addressing community health led to improvements in individual-level health outcomes or system-level reform. In the three initiatives in which coordination was confused with collaboration and stakeholders directed their efforts toward coordination rather than collaboration, they were unable to show improvement in health outcomes. Perhaps the confusion about collaboration and its impacts on community health can be reduced by examining the common features present in highly successful collaboratives rather than simply if collaboration, broadly speaking, was employed.

**Process Quality: A Key Feature of Effective Collaboration**

Chrislip and Larson (1994) have identified the features common to highly successful collaborative initiatives. Working from a 3-year qualitative and quantitative analysis of 52 cases of highly successful collaboration identified by the National Civic League—with success being measured by the ability of the collaborative to demonstrate tangible results—they discovered two features present in all 52 cases. The first is strong process leadership.
The second, and the focus of this article, is the presence of an open and credible process.

Strong process leadership signals a shift away from a hierarchical model in which strong leadership entails strong advocacy of a particular point of view. Process leadership, instead, involves bringing the appropriate people to the table and keeping them there through difficult periods, facilitating the expression of divergent points of view in a manner that respects difference yet fosters convergence and making sure that all stakeholders feel competent, trusted, and valued throughout the process (Chrislip and Larson 1994, 53). In a literature review of collaborative leadership research in public health, Larson, Sweeny et al. (2002) reported 11 studies, none of which were reported in Roussos and Fawcett (2000) and Kreuter et al. (2000), that found a strong relationship between the presence of strong process leadership and improvements in health outcomes ranging from the decline of infant mortality rates to improvements in cardiovascular health (Larson et al. 2002, 8–12).

The second factor associated with all 52 cases of successful collaboration is the presence of an “open and credible process” (Chrislip and Larson 1994, 53). This factor refers to the extent that stakeholders perceive the process to be fair and authentic. That is, stakeholders perceive that all are treated equally and secure in that decisions have not already been made in advance with the process simply serving as legitimation for those decisions. Stakeholders must be confident that the process is free from behind-the-scenes manipulation and that safeguards are in place to check the disproportionate influence of powerful individuals. In short, stakeholders must feel secure that all involved in the process have equal opportunity to directly influence the decision made and that decision is likely to have some impact on the root problem the participants are addressing. An open and credible process embodies what the procedural justice literature refers to as the “voice effect” (Lind et al. 1990; Tyler et al. 1997, 88–90): the strong tendency for people to see processes as more fair if they have an opportunity to influence the process before decisions are made.

One of the fundamental insights of the research on process quality is that processes have a communicative function (Tyler and Lind 1992; Lind 2001). When people are invited to participate in collaborative activity, they must make a quick, often intuitive, judgment as to the likelihood that they will be exploited or rejected by the others. Although there may be significant advantages of collaboration, a judgment that there is a relatively high likelihood of either exploitation or rejection will lead people to pursue lower risk, lower reward self-oriented goals. The most common cognitive device that people use to determine the risks inherent in collaborative activity is their impressions of
the fairness of procedures. The structural features of a process and the quality of treatment that others extend to participants in the process are among the first available, easiest to ascertain, and most reliable indicators of one’s security in a social situation (it is important to note that the ways others are treated, as well as the way one’s self is treated, gives important information about the quality of a process). Fairness judgments thereby “serve as a proxy for interpersonal trust in guiding decisions about whether to behave in a cooperative fashion” (Lind 2001, 56). People continually read processes for relational cues that they can use to ascertain their status within a group. When people perceive that they are being treated fairly—understood in terms of positive attributions of trust, neutrality, and standing—they will, in turn, feel valued, respected, and cared for by the group. The result is that they will come to see their individual identity in terms of their group membership: an identification that, in turn, results in an increased commitment to the groups’ projects and goals. Moreover, they will come to see the integrity of the process as an expression of the integrity of the group itself. Hence, the “structural” and “relational” aspects of procedural fairness are mutually constitutive, with deficits in one or more of the structural conditions leading to negative relational judgments, and negative relational judgments leading to the belief that one or more of the structural conditions are being manipulated.

We have examined the impacts of stakeholders’ perception of process quality on collaborative activity in two field studies. The first study resulted from complaints on the part of members of quasi-governmental structures known as Metropolitan Planning Organizations. These organizations were created by the U.S. Congress to engage in long-range strategic planning on transportation and other issues. Most of us know these organizations as “regional councils of government.” The complaints about how transportation dollars were allocated among the member governments resulted in the construction of a scale that allowed participants in these intergovernmental collaboratives to rate the quality of the processes that they employed to make resource allocation decisions. The results disclosed a striking correspondence between the rated quality of the processes and the effectiveness of the long-range strategic planning, as assessed by the member governments (Goetz et al. 2002).

The second field study involved a 10-year follow-up to the Colorado Healthy Communities Initiative (CHCI; Larson et al. 2002). The Colorado Trust’s CHCI was begun in 1992 to assist communities in defining their own vision of a healthy community and working to achieve that vision. The $8.8 million initiative was modeled on healthy city and healthy community programs developed by the World Health Organization. The 28 communities completed a variety of projects consistent with each community’s
vision of a healthy community. Approximately 10 years after the completion of the projects, a follow-up study was commissioned to evaluate the success of these projects, what factors accounted for the success of the projects, which projects were still active, and what accounted for the sustainability of the projects that remained active. When the active sites were compared with the inactive sites, large differences were found. Specifically, the active sites received ratings on the “process quality scale” significantly higher than those received by the inactive sites (5.05 versus 3.43 on a scale of 1 to 6, \( p < .001 \), \( t \) for independent samples). The correspondence between success and sustainability was so close that it is reasonable to assume that the quality of the process underlies both.

The opportunity for a more direct test of the relationship between process quality and programmatic success came with the evaluation of the Colorado Nurse–Family Partnership.

The Context of the Present Study:
The Nurse-Family Partnership (NFP)

The NFP is a program in which specially qualified nurses visit first-time, low-income single mothers at least biweekly from as early in pregnancy as possible, through the child’s second birthday. Under this voluntary program, nurses work with mothers on improving their prenatal health, understanding child development and using that knowledge to foster effective parenting strategies, and becoming self-sufficient. The NFP has been rigorously tested through ongoing, longitudinal randomized trials in Elmira, New York, Memphis, and Denver. In the 15-year follow-up of the Elmira sample (Olds and Korfmacher 1998), the program produced results including a 79% reduction in child abuse and neglect, a 56% reduction in children’s emergency room visits when injuries were detected, a 30-month reduction in mother’s use of welfare, and 69% fewer arrests for the mothers during the 15 years after delivery of first child. The nurse-visited children had 54% fewer arrests compared to their control-group counterparts.

The NFP is being replicated nationally by the NFP national office, which is devoted to careful reproduction of the program in local communities to increase the likelihood that the program will continue to produce beneficial effects for mothers and children consistent with those found in the scientifically controlled trials. The national NFP spends considerable effort ensuring that local communities and organizations are ready to conduct the program, provides intensive training and technical support to nurses hired by
local organizations to conduct the program, and provides a Web-based information system into which every local organization enters data on program implementation and maternal and child outcomes that are benchmarked against standards achieved in the trials. Data are entered into a “Clinical Information System” that is used as a foundation for continuous quality improvement of program operations and outcomes.

The Colorado NFP is unique in that it devotes greater attention than other states in the use of a collaborative site-development strategy. In Colorado the adoption of the NFP is done through collaborative partnerships including county health departments, community health agencies, county human services departments, school boards, local head starts, county commissioners, and business and civic leaders. The program is then implemented by local health agencies and supported by and planned for by these collaborative partnerships. At the time of this study, 16 collaborative partnerships, ranging from citywide to multicounty collaboratives, were in place, with more than 2,800 families in 50 of Colorado’s 64 counties enrolled in the NFP (presently there are 18 sites operative in Colorado). These collaborative partnerships are facilitated by Invest in Kids (IIK), a nonprofit organization based in Denver, Colorado that partners with communities to implement evidence-based programs targeting children (prenatal to age 5), particularly those from low-income families. IIK has played a critical role in bridging the gap between the resources and commitments of the community and the requirements of implementing the NFP. IIK is most heavily involved in the formative stages of the collaborative partnership: recruiting stakeholders from key sectors, building community commitment to establish the program, and facilitating the implementation of the program by a community agency. IIK also facilitates the establishment of a community advisory board in each site that provides ongoing support for the program, including facilitating interagency collaboration between multiple sectors in areas such as funding needs, client referral, and hiring home-visitation nurses. In addition, IIK provides ongoing technical assistance and support to the 16 communities, acting as a bridge between the communities and the National NFP office.

We have evaluated the site development and implementation strategies of IIK and these collaborative partnerships since 1999. The primary interest is in determining the relationships between features of community collaboration, particularly the quality of the process and the program outcomes reported by the Colorado NFP. This evaluation has found consistent relationships between process quality and program outcomes (Hicks and Larson 2003). These program outcomes are of two kinds. One is mother attrition. The second involves those variables associated with birth: birth weight, the number of premature births, and subsequent pregnancies. Our
evaluations of the program have consistently found relationship between the early community collaboration processes and both sets of outcomes. Of course, it reasonable to assume that these two outcomes are themselves correlated: Staying in the program, which, of course, would result in a higher “program dose,” ought to be associated with higher program outcomes. The issue is whether or not the early community collaboration processes predict both sets of outcomes when variations among individuals are controlled for. For this reason, we undertook hierarchical linear modeling (HLM) analysis of these two sets of program outcomes. Unfortunately, our analysis of birth-related outcomes failed to produce any significant predictive relationship between process quality and those birth-related outcomes.

Our analysis did find, however, that process quality—measured by the Process Quality (PQ) scale described below—is associated with the first program outcome, mother attrition. Specifically, sites have less attrition as stakeholders perceive the process as free from bias. The present research is motivated by the need to further refine and test this finding.

Nationally and in Colorado, one of the most significant challenges to the replication of the NFP has been that participants are dropping out of the program at higher rates than families did in the scientifically controlled trials. Considerable attention has been devoted to identifying individual participant characteristics that predict attrition and nurse characteristics (Moritz et al. 2003). Moreover, analyses of data from the national Clinical Information System show that there are significant differences among sites in the rates of attrition, even after controlling for individual characteristics (McClatchey et al. 2001). This same variability among sites occurs in Colorado, with 16 sites ranging from attrition rates of 0.28 to 0.57 with no major clusters of sites. It is not clear what it is about sites that accounts for the intersite variation in participant attrition; each of the sites receives equivalent amounts of funding, does not vary significantly in nurse or client characteristics, and receives equivalent implementation support from the National NFP office and IIK. The sites do, however, vary in terms of the quality of the collaborative process.

To investigate possible multilevel effects on attrition involving individual and site-level variables, we tested the following two hypotheses:

**Hypothesis 1:** Attrition will be related to individual participant-level variables shown to predict attrition in previous studies on the NFP as well as with the degree to which sites were developed according to principles of high-quality collaborative process.

**Hypothesis 2:** Collaborative processes involved in the development of the sites will account for a significant proportion of the variance in attrition beyond that which is accounted for by individual participant-level variables.
Method

Scope of the Study

We analyzed data for all 16 site locations in Colorado in which the NFP had been implemented between 1999 and 2003. These 16 sites averaged 176.5 participants. The sites ranged in size from 51 to 294 participants, for a total 2,824 participants.

Data Collection

Site-level data collection. We collected site level data through interviews that included descriptions and ratings of the process, levels of stakeholder commitment to the process, community support, and history of collaboration in the community. The first two authors conducted the interviews. At the time of the interview, stakeholders also completed the Working Together instrument (Chrislip and Larson 1994), which measures motivation and interagency collaboration, and the PQ Scale, which measures the perceived fairness and authenticity of the collaborative process. These stakeholders were identified by IIK as members of the initial working group. The number of stakeholders in each site varied with the site development strategy employed in that community, ranging from 5 to 20 stakeholders in each site, resulting in a total of 186 respondents. The participation rate was 97%, with 5 identified stakeholders saying that they were not sufficiently involved in the early process to respond to the interview. Site-level scores were calculated by simply averaging the ratings for all stakeholders associated with a particular site. The number of stakeholders in each site should not be read as indicating anything about the openness (or closed nature) of the collaborative process; rather, coalition size varies considerably because of the geographic or population characteristics of the sites. The site-level data assessed the process employed by the early stakeholder group in bringing the program to the community. The data was collected after the stakeholder groups had decided to bring the program to their communities. This data was collected approximately 4 years prior to the program attrition data to allow ample time for the program to be implemented and for its pattern of success to become established.

Participant-level data collection. Participant-level data nested within sites are continuously collected by the visiting nurses, entered into the Web-based Clinical Information System, and organized and analyzed by biostatisticians from the National Center for Children, Families, and Communities, in the
School of Nursing at the University of Colorado in Denver, which evaluates the nationwide implementation of the NFP. The nurses collected all participant-level data. Descriptive statistics for the participant-level variables are shown in Table 1.

**Table 1**

**Description of Participant-Level Variables**

<table>
<thead>
<tr>
<th>Participant-Level Variables</th>
<th>Description and Scale of Participant-Level Variables (Level 1)</th>
<th>( M )</th>
<th>( SD )</th>
</tr>
</thead>
<tbody>
<tr>
<td>Dropout</td>
<td>Outcome variable of participant status; dichotomous (1 = dropped, 0 = active)</td>
<td>0.46</td>
<td>0.50</td>
</tr>
<tr>
<td>Education</td>
<td>The number of years of education completed upon entering the program; continuous</td>
<td>10.97</td>
<td>2.38</td>
</tr>
<tr>
<td>Nurse visits</td>
<td>Number of nurse visits during the participant’s pregnancy; continuous</td>
<td>8.75</td>
<td>4.26</td>
</tr>
<tr>
<td>Mastery</td>
<td>( z ) score for mastery of program content; continuous</td>
<td>100.51</td>
<td>10.30</td>
</tr>
<tr>
<td>Nurse attrition</td>
<td>Work status of nurse; dichotomous (1 = nurse no longer working, 0 = nurse still working)</td>
<td>0.12</td>
<td>0.32</td>
</tr>
<tr>
<td>Race</td>
<td>Participant’s race; dichotomous (1 = African American, 0 = other)</td>
<td>0.02</td>
<td>0.13</td>
</tr>
<tr>
<td>Working</td>
<td>Participant’s work status; dichotomous (1 = currently working, 0 = not working)</td>
<td>0.37</td>
<td>0.48</td>
</tr>
<tr>
<td>Age</td>
<td>Participant’s age at first completed visit; continuous</td>
<td>19.80</td>
<td>4.32</td>
</tr>
<tr>
<td>Mental score</td>
<td>( z ) score mental health; continuous</td>
<td>100.51</td>
<td>9.70</td>
</tr>
</tbody>
</table>

Overview of the Site-Level Measurement Instrument

The PQ Scale developed by Hicks and Larson is a 15-item questionnaire addressing various aspects of process quality. The questionnaire format uses a 6-point scale with the following responses: 1 = *strongly disagree*, 2 = *disagree*, 3 = *disagree more than agree*, 4 = *agree more than disagree*, 5 = *agree*, 6 = *strongly agree*. Items 3, 8, and 9 were reverse scored, because higher agreement with the actual questionnaire statement implied a negative description of process quality.

An exploratory factor analysis (EFA) was conducted to determine interpretable subscales of the PQ questionnaire. Principle components analysis (PCA) using varimax rotation was used to determine the number of factors...
present and the factor structure of each of the identified factors. The factor structure was identified by items with factor loadings greater than 0.30. The final solution consisted of three factors. Factor 1, labeled structural integrity, consists of items that deal with the establishment and consistent application of fair rules of distribution. This factor encompasses what Leventhal (1980) identified as the structural features of procedural fairness. Factor 2, labeled authenticity, consists of items that deal with the openness and credibility of the process, including whether stakeholders perceive the process as free from undue influence from special interests outside the process, whether the process itself can generate binding decisions rather than simply confirming decisions already made, and whether the stakeholders perceive that they have equal standing. This factor encompasses what Chrislip and Larson (1994) referred to as the presence of an “open and credible process” and Tyler and Lind (1992, 2001) as the relational features of procedural fairness. Factor 3, labeled equity, consists of items that deal with the actual distribution of outcomes, whereby structural integrity refers to the procedural criteria for distribution—specifically, if the process distributes outcomes equally and does not discriminate on the basis of organizational affiliation. This factor encompasses what Thibaut and Walker (1975) refer to as “distributive justice.”

The results of the exploratory analysis of site-level variables led to the identification of the PQ subscale (Authenticity) as the site-level predictor to be included in the final multilevel analysis. The full scale was not used because in the exploratory analysis of site-level variables, the full-scale score and the other factor scores were not significantly associated with attrition. Descriptive statistics for the Process Authenticity Scale are shown in Table 2.

The observed reliability of the total PQ Scale has been found to average 0.87 for all studies using this measure. The reliability of the PQ Authenticity subscale is 0.62 for all studies using the measure, but the reliability is 0.55 for the present study. In our initial analysis of the relationships between stakeholder processes and subsequent attrition, we included all three of the subscale scores and the total scale score. Although the authenticity subscale has low reliability, it is the subscale that had the closest association with attrition and is the primary process quality measure retained for the HLM analysis.

**Multilevel Data Analysis Approach**

HLM is a method that can be used to answer questions involving site-level predictors of person-level outcomes. In this study, the person-level outcome of interest is the participant attrition rate. The site-level predictor
is a variable of process quality. The multi-level model used in this study was the hierarchical generalized linear model (HGLM), which is used when the outcome of interest is dichotomous. The outcome variable attrition was collected as a dichotomous variable in which 1 = dropped and 0 = active.

Results

Results of the Unconditional Model

To gauge the magnitude of variation between sites in the outcome variable attrition, a model with no predictors at Level 1 and no predictors at Level 2 was estimated. The typical probability of dropping out, calculated from the intercept coefficient, was 0.45, which was slightly lower than the mean calculated from the descriptive statistics of the attrition variable for which the mean was calculated as 0.46.

Results of the Multilevel Model

To account for additional variation in participant attrition, we estimated a conditional model with four (4) identified participant level (Level 1)

### Table 2

**Description of Site-Level Variables**

<table>
<thead>
<tr>
<th>Site-Level Variables</th>
<th>Description and Scale of Site-Level Variables (Level 2)</th>
<th>M</th>
<th>SD</th>
</tr>
</thead>
<tbody>
<tr>
<td>PQ Authenticity subscale</td>
<td>Process Quality (PQ) Questionnaire (Factor 2: Authenticity Subscale) consisting of the sum of the responses for Items 3, 8, and 9; continuous</td>
<td>12.34</td>
<td>1.67</td>
</tr>
<tr>
<td>PQ Scale (Item 3)</td>
<td>“Often decisions are made in advance and simply confirmed by the process”; 6-point rating scale</td>
<td>3.82</td>
<td>0.73</td>
</tr>
<tr>
<td>PQ Scale (Item 8)</td>
<td>“In the process, some people’s ‘merits’ are taken for granted while other people are asked to justify themselves”; 6-point rating scale</td>
<td>4.37</td>
<td>0.64</td>
</tr>
<tr>
<td>PQ Scale (Item 9)</td>
<td>“In the process, strings are being pulled from the outside, which influence important decisions”; 6-point rating scale</td>
<td>4.15</td>
<td>0.75</td>
</tr>
</tbody>
</table>
predictors of attrition and one (1) site-level (Level 2) predictor. PQ Authenticity subscale was the variable included at the site level (Level 2). Using the nonlinear Bernoulli model, the logit link function, and the unit-specific model described in Equations 1 and 2 (below), 2,824 Level 1 participants \( (n = 2,824) \) were used to estimate the conditional Bernoulli model for Level 2 along with 16 Level 2 units \( (n = 16 \text{ program sites}) \).

**Level 1 Model:** \[
\text{Prob}(Y = 1|\beta) = P \\
\log \left( \frac{P}{1-P} \right) = \beta_0 + (\beta_1 \times \text{Nurse Visits}) + (\beta_2 \times \text{Mastery}) + (\beta_3 \times \text{Education}) + (\beta_4 \times \text{Nurse Attrition}) + r_{ij} \\
(1)
\]

**Level-2 Model:**
\[
\begin{align*}
\beta_0 &= \gamma_{00} + (\gamma_{01} \times \text{PQ Authenticity subscale}) + u_0 \\
\beta_1 &= \gamma_{10} + u_1 \\
\beta_2 &= \gamma_{20} + u_2 \\
\beta_3 &= \gamma_{30} + u_3 \\
\beta_4 &= \gamma_{40} + (\gamma_{41} \times \text{PQ Authenticity subscale}) + u_4 \\
(2)
\end{align*}
\]

Table 3 identifies the results for the multilevel model for the outcome attrition. The table includes log-odds for the intercept as well as the four independent predictor variables: the number of visits during pregnancy (nurse visits), the mastery score on entry in the program (mastery), the education level of the participant on entry in the program (education), and whether or not the nurse left the program (nurse attrition). Table 3 also includes the results for the PQ Authenticity subscale at the site level.

As a result of including nurse visits, mastery, education, and nurse attrition within the Level 1 model, these variables accounted for 10.2% of the variation in the outcome attrition. As a result of including nurse visits, mastery, education, and nurse attrition at Level 1 and adding the PQ Authenticity subscale at the site level (Level 2), the final multi-level HGLM model accounted for 22.8% of the variation in the outcome attrition, for an increase of 12.6% additional variance explained by the multilevel model.

We see here that number of visits during pregnancy (nurse visits) is associated with lower log-odds of dropping out \( (\gamma_{10} = -0.117, t = -10.05, p < .05) \), mastery scores (mastery) are associated with higher log-odds of dropping out \( (\gamma_{20} = 0.010, t = 2.19, p < .05) \), the number of years of education (education) is associated with lower log-odds of dropping out \( (\gamma_{30} = -0.058, t = -2.63, p < .05) \), and nurse attrition (nurse attrition) is associated with higher log-odds of dropping out \( (\gamma_{40} = 2.434, t = 6.72, p < .05) \) controlling for all other variables. The assumption of linearity of the logit was tested in...
accordance with the Box–Tidwell approach (Hosmer and Lemeshow 1989, as found in Tabachnick and Fidell, 2001). The nurse-visit variable was the only variable in violation of the assumption. An alternative model was created in which “nurse visits” was replaced by its logarithm transform; however, the alternative model did not differ substantively from the original model with the untransformed value. Therefore, the simpler, untransformed model was analyzed as the final model for this article.

At the site level (Level 2), the log-odds of dropping out was significantly related to the PQ Authenticity subscale. The log-odds of dropping out was inversely related to the PQ Authenticity subscale holding other predictors constant ($\gamma_{01} = -0.188, t = -4.37, p < .05$). A 1-unit increase in the PQ Authenticity subscale reduces the log-odds of dropping out by 0.188. Associated with this is a relative odds ratio of $\exp\{-0.188\} = 0.83$. That is, if we compare 2 participants who are similar in other ways but differ by 1 unit in the PQ Authenticity subscale, we can expect the odds of dropping out of the higher PQ Authenticity subscale participant to be 0.83 times (less than 1 or even odds) the odds of dropping out of the lower PQ Authenticity subscale participant.

In analyzing the direct effect of PQ Authenticity on dropout, as the PQ Authenticity subscale increases, the probability of dropping out decreases, as shown in Table 4.

The PQ Authenticity subscale, a three-item subscale, with items scaled 1 to 6, may range from a low of 3 to a high of 18. Therefore, a change in the PQ Authenticity subscale of six units represents an average change in the three items of two units each. Thus, changing the perception of the process from slightly inauthentic to authentic would be associated with a 24.2% decrease in dropout.

### Table 3

<table>
<thead>
<tr>
<th>Fixed Effects</th>
<th>Coefficient</th>
<th>SE</th>
<th>t Ratio</th>
</tr>
</thead>
<tbody>
<tr>
<td>Intercept, $\gamma_{00}$</td>
<td>-0.187</td>
<td>0.075</td>
<td>-2.51*</td>
</tr>
<tr>
<td>PQ Authenticity subscale, $\gamma_{01}$</td>
<td>-0.188</td>
<td>0.043</td>
<td>-4.37*</td>
</tr>
<tr>
<td>Nurse Visits slope (B1), $\gamma_{10}$</td>
<td>-0.117</td>
<td>0.012</td>
<td>-10.05*</td>
</tr>
<tr>
<td>Mastery slope (B2), $\gamma_{20}$</td>
<td>0.010</td>
<td>0.004</td>
<td>2.19*</td>
</tr>
<tr>
<td>Education slope (B3), $\gamma_{30}$</td>
<td>-0.058</td>
<td>0.022</td>
<td>-2.63*</td>
</tr>
<tr>
<td>Nurse Attrition slope (B4), $\gamma_{40}$</td>
<td>2.434</td>
<td>0.362</td>
<td>6.72*</td>
</tr>
<tr>
<td>PQ Authenticity subscale, $\gamma_{41}$</td>
<td>-0.398</td>
<td>0.174</td>
<td>-2.29*</td>
</tr>
</tbody>
</table>

*Significant at $p < .05$. 

Hicks et al. / Colorado Nurse–Family Partnership 467

---

**Downloaded from http://erx.sagepub.com by guest on February 17, 2009**
Also at the site level (Level 2), the slope of nurse attrition was significantly related to the PQ Authenticity subscale. The slope of nurse attrition was inversely related to the PQ Authenticity subscale holding other variables constant. As scores on the PQ Authenticity subscale increase, the impact of nurse attrition on participant dropout decreases ($\gamma_{41} = -0.398$, $t = -2.29$, $p < .05$). If a nurse terminates employment, the probability of the mother (i.e., client of the nurse) dropping out is 90.4%; however, as scores on the PQ Authenticity scale increase, the probability of dropping out continues to decrease even if nurses terminate employment.

In analyzing the direct effect (described previously) and the indirect effect (described by reducing the impact of nurse attrition) of PQ Authenticity on dropout, as scores on the PQ Authenticity subscale increase, the probability of dropping out decreases from the baseline of 90.4% (within the context of nurses terminating employment), as Table 5 shows.

### Table 4
**Percentage Change in Probability of Attrition, Direct Effects of PQ**

<table>
<thead>
<tr>
<th>PQ Authenticity Subscale</th>
<th>% Change in Probability of Attrition</th>
</tr>
</thead>
<tbody>
<tr>
<td>Increase PQ Authenticity subscale by:</td>
<td></td>
</tr>
<tr>
<td>1-unit increase</td>
<td>–4.6%</td>
</tr>
<tr>
<td>2-unit increase</td>
<td>–9.1%</td>
</tr>
<tr>
<td>3-unit increase</td>
<td>–13.3%</td>
</tr>
<tr>
<td>4-unit increase</td>
<td>–17.2%</td>
</tr>
<tr>
<td>5-unit increase</td>
<td>–20.9%</td>
</tr>
<tr>
<td>6-unit increase</td>
<td>–24.2%</td>
</tr>
</tbody>
</table>

### Table 5
**Percentage Change in Probability of Attrition, Indirect Effect of PQ**

<table>
<thead>
<tr>
<th>PQ Authenticity Subscale</th>
<th>% Change in Probability of Attrition</th>
</tr>
</thead>
<tbody>
<tr>
<td>Increase PQ Authenticity subscale by:</td>
<td></td>
</tr>
<tr>
<td>1-unit increase</td>
<td>–4.1%</td>
</tr>
<tr>
<td>2-unit increase</td>
<td>–9.5%</td>
</tr>
<tr>
<td>3-unit increase</td>
<td>–16.3%</td>
</tr>
<tr>
<td>4-unit increase</td>
<td>–24.7%</td>
</tr>
<tr>
<td>5-unit increase</td>
<td>–34.1%</td>
</tr>
<tr>
<td>6-unit increase</td>
<td>–44.0%</td>
</tr>
</tbody>
</table>
As a result, changing the perception of the process from slightly inauthentic to authentic would be associated with a 44.0% decrease in dropout within the context of nurses terminating employment.

**Discussion**

We hypothesized that the quality of the process of building community collaboration would account for a significant proportion of the variance in attrition beyond that which is accounted for by characteristics of individual participants. This hypothesis was supported.

The most striking thing about the results is not simply that the hypothesis is supported but rather that it is the strength of the association found between process quality and participant attrition. We controlled for the individual participant variables found to be most strongly associated with attrition, as identified by the program developer and research staff at the National Center for Children and Families, and we added a variable to the Level 1 model (nurse attrition) because the field staff saw so clearly its relationship to attrition and advocated its inclusion in the Level 1 model. Yet even with the best Level 1 model that we could construct, we still find that model accounting for 10.2% of the variance in attrition. The Level 2 model—based entirely on what we might call the authenticity of the communication that occurs within the collaborative initiative—accounted for an additional 12.6% of the variance in mother attrition. Furthermore, when we factor in the relationship between nurse and mother attrition, we obtain even more insight into the influence of process quality on attrition. That is, within the context of nurse attrition, lower levels of process quality are associated with greater impact of nurse attrition on mother attrition. We believe that lower levels of process quality may heighten the sensitivity of mothers to the more subtle but important aspects of the process surrounding the program and, therefore, increase the likelihood that the mother will leave if the nurse leaves.

This finding supports and may inform research on the relationship between coalition functioning and effective program implementation. Kegler and colleagues (1998) identified several factors that facilitated and obstructed effective program implementation in the 10 Project ASSIST coalitions in North Carolina. They found that those coalitions in which the stakeholders themselves generated the vision guiding program implementation “had higher levels of implementation than coalitions that were lacking a local vision’’ (p. 231). Furthermore, they found that “coalitions with minimal member input had lower levels of implementation’’ (Kegler et al. 1998, 234). Both of these factors—member
input and source of vision—are intimately related to stakeholder perception of process authenticity. Stakeholders perceive a process as authentic when their input is valued and they believe that the collaborative itself, rather than some external agency, has the authority to generate plans and have those plans guide decision making. Feinberg et al. (2004) found that the internal functioning of coalitions was linked to the perceived effectiveness of community prevention coalitions funded by the Communities That Care (CTC) initiative in Pennsylvania. They found that community readiness to foster a preventive coalition was strongly associated with perceived effectiveness and that, moreover, the effect of community readiness on effectiveness was mediated through the quality of the coalition’s internal functioning. Internal functioning refers to both the motivations of coalition members and the quality of coalition functioning as measured by such features as the time and energy that coalition members were willing to spend on implementation efforts, the sense that the coalition was moving in a positive direction, the personal benefits acquired through coalition membership, and the clarity of the implementation plan. There may be an important relationship between what Feinberg and his colleagues understand as internal coalition functioning and board efficacy and what we have described as process quality.

One additional fact makes the finding in the present research even more surprising: Site-level data on process quality was collected approximately 4 years before the programmatic data on attrition were collected. That is, the process quality data were gathered from interviews and scaled responses of community stakeholders who were involved in bringing the NFP to their community and were developing a collaborative initiative to achieve this result. Four years later, when the program had been implemented and a cohort of young mothers had completed or was completing the program, the attrition data were summarized for each of the 16 programs. Those early stakeholder groups that employed an authentic process were able to plan and implement the program in such a way that it had significantly lower attrition rates. Stakeholders’ accounts of the collaborative process gathered in our interviews, our prior field studies examining metropolitan planning organizations (Goetz et al. 2002) and the CHCI (Larson et al. 2002) as well as the substantial body of research on the fair process effect cited above lead us to believe that if the early stakeholder group devoted its energies to competing for resources, building its back home agencies, expanding spheres of influence, or pursuing individual agendas the authenticity of the process would have suffered and the lack of confidence in that process would have had ripple effects that ultimately impacted the willingness of others to commit their time and energy to implementing the decisions of this early stakeholder group.
That these effects will ripple for years may seem difficult to accept. Yet in the CHCI (Larson et al. 2002), a strong relationship between process quality and success and sustainability of community initiatives was discovered 10 years after the initiatives began. Gomez et al.'s (2005) longitudinal study of the Communities That Care Coalitions (CTC) in Pennsylvania also found that coalition sustainability was related to early coalition function. Both of these studies suggest that the perceived quality of the collaborative partnership in its formative stages is related to the success and sustainability of the collaborative partnership several years later. How can we understand the perseverance of these effects? Let us offer our explanation.

Tyler and Bladder (2000) convincingly demonstrated that perceptions of procedural fairness are the primary motivation for voluntary, discretionary cooperation within workplace settings. Conversely, they show that the perception of unfairness is the primary motivation for antisocial, combative resistance. Applying this insight to collaboration, if stakeholders perceive the process as unfair, they may be more likely to abandon collaboratives; or worse, they may remain and find ways to manipulate the process to garner more resources for themselves at the expense of others. They will not, however, see their actions as unethical but as the natural outcome of the process itself. The result is a vicious circle of selfishness and the eventual collapse of the collaborative process. On the other hand, when stakeholders perceive the process as fair they will act cooperatively even when they receive less than what they hoped for. And they will take others’ needs and desires into consideration in forming their own convictions. The result is a virtuous circle whereby the initial energy invested into the collaborative fosters greater commitment to the process and stakeholders continue to rededicate themselves and their resources to sustaining joint initiatives.

This sense of real cooperation is crucial because within civic collaborations, when people participate, they must make a quick, often intuitive judgment as to the likelihood that they will be exploited by others (Tyler and Lind 1992, 2001; Lind 2001). Although there may be significant advantages of cooperation toward group goals, a judgment that there is a relatively high likelihood of exploitation will lead people to pursue lower risk, lower reward self-oriented goals. However, a high level of cooperation is possible even with self-oriented individuals, but they require trust in others. Cooperation and altruistic behavior is a rational response in groups with unselfish interests and this behavior can be passed-on between people (Skryms 1996; Katz and Shapiro 1985; Ostrom 2000).

Cooperation and commitment, like other patterns of authentic behavior, are contagious. Behaviors signaling both high cooperation and low cooperation...
are transferable despite a high turnover of members. This has been confirmed in the research on teamwork. LaFasto and Larson (2001) have reported research with 6,000 team members on the importance of creating and sustaining mental, physical, and emotional energies in teams that are unusually successful. These energies often spread through teams to produce outcomes that go beyond the performance expectations one would have of individual team members.

We believe that this happened with IIK. The results are very clear empirical relationships between the commitment of the stakeholder groups and the dropout rates of the young mothers in the program several years later. We call this fundamental pattern of a learned level of cooperation among people the transfer of commitment.

This transfer begins with the role played by IIK in developing the sites. These sites vary considerably in terms of the circumstances and conditions encountered by IIK in the early stages of site development. IIK site development strategies were responsive to the communities’ historically established patterns of working together on social issues; therefore, the site development strategies used to establish the collaborative dedicated to deliberating over the adoption and implementation of the NFP varied considerably from community to community. Because of the extant differences among communities, the resulting processes varied considerably in terms of their overall quality as perceived by the participants.

What emerged as consistent across communities, however, was the perception of IIK as genuinely concerned about improving the future for children, as advocating the well-being of children, as investing their time and resources in children, and in impacting in real and tangible ways the welfare of children. They are a group of private citizens who have no vested interest beyond their concern for improving the quality of children’s lives. They bring a problem focus to the community and promote and sustain a cooperative process that is targeted on the problem. The model of cooperative behavior that is created by their commitment is transferred to the stakeholder group (the community). This happens in at least three ways. First, IIK recruits stakeholders who have prior experience in community collaboratives and are committed to using collaborative processes to implement the NFP. Second, IIK continuously communicates with the stakeholders, providing proactive responses to any emerging problems. Third, IIK facilitates the transition of the community coalition into an advisory board that provides support to the implementing agency, making concerted efforts to retain key individuals who are deeply committed to the program on the advisory board. The commitment of the community coalition may then be transferred to the nurse supervisor.
via her or his participation on the advisory board. Moreover, the commitment of implementing agency representatives, including the nurse supervisor, may contribute to the building of an organizational climate that fosters the alliance between the nurse home visitor and the client, a connection that is highlighted in the research on the relationship between organizational climate and the fostering of positive relationships between employees and clients (Schneider, White, and Paul 1998) and the development of the therapeutic alliance between mental health providers and clients (Glisson 2002; Watson and Greenberg 1994). If the nurse supervisors are then able to transfer their commitment to the home visitation teams, they in turn may transfer this commitment to the client group (the first-time mothers). We also found that process quality mediates the impact of nurse attrition on mother attrition. One explanation is that when process quality is lower, the nurses’ decision to leave the NFP may be perceived as more reflective of the program itself and, thus, may influence the mother’s decision to commit to continued participation. The most direct relationship to program outcomes is, then, those clients who keep their appointments and stay engaged in the program.

It took many years and much hard work for health care professionals and social scientists to conclude that events and experiences in infancy and early childhood had lasting effects on the child’s, and then later, the adult’s health and well-being. In fact, this assumption underwrites both the theoretical foundations of the NFP and the body of evidence demonstrating its success. We are arguing analogously that the events and experiences in the infancy and early stages of the collaborative process have lasting effects on the commitment to and ultimate success of the programs that process engenders.

Limitations

Field studies are frequently limited in external validity, so generalizations must be made cautiously. This study is limited in the sample size for the Level 2 variables, as the Level 2 model is based on 16 program sites. Even though these 16 sites involve all of the sites within Colorado and serve 2,824 young mothers, we are still restricted to 16 PQ Authenticity subscale scores in our Level 2 analysis. Sixteen is a small sample for a study that is correlational in nature.

The PQ Authenticity subscale has a relatively low reliability. Reliability coefficients were calculated for the overall PQ questionnaire as well as for the individual subscale used in the multi-level analysis (Hintz, 2000). The reliability of the overall questionnaire for this study was found to equal 0.85 (Cronbach’s $\alpha$) and reliability of the three-item Process Bias subscale was
found to equal 0.55 (Cronbach’s $\alpha$). Because reliability is a function of the number of items, among other things, this reliability for a three-item sub-scale is not unexpected. But low reliability is a limitation even though the subscale seems to be an unusually accurate predictor of attrition.

**Future Directions**

One needed improvement in the line of research that we are pursuing involves sample size. Whenever predictions are made from site-specific variables, sample size becomes a problem. We are addressing this problem in two ways. First, we are in the third year of a longitudinal study on a program designed to increase the social competence and to decrease aggressive and problematic behavior in young children. This program has already exceeded the number of NFP sites that we investigated and promises to provide 4 to 5 times as many sites. Second, we are using the data from the NFP research to engage in agent-based modeling, a process that allows us to replicate the essential features of the sites that we studied, to modify some of the qualities of collaboration that we measured, and to examine the trends in collaborative behavior produced by these modifications. The modeling will then be used to provide insight into the dynamics of the transfer of commitment. Agent-based modeling allows us to explore patterns that emerge from the data collected from the participants of the collaborative activities of the 16 sites that we examined in the research reported here.

**Conclusion**

We presented the problem in the introduction of the continuing attempt to assess whether or not, and to what extent, collaboration promotes community health. We are now, on the basis of our findings, in a position to offer the following conclusions. First, there is a clear and identifiable relationship between the quality of the collaborative process and community health program success, at least as these were operationalized in this research. Second, increased clarity about some very subtle relationships that may span years or even a decade can be achieved through greater explication of this broad construct called collaboration and what specific features that we are attempting to measure. The present research suggests that an important and potentially predictive feature of collaboration, and one that we have the means to measure, is process quality and, more specifically, the authenticity of that process.
References


Darrin Hicks, PhD, is an associate professor in the Department of Human Communication Studies at the University of Denver. He teaches and conducts research in public deliberation, argumentation theory, and community collaboration. He has recently published articles on public deliberation in The Quarterly Journal of Speech, Rhetoric and Public Affairs, and Argumentation. He is currently working on a book, The Other Side of Reason, which examines the evolution of standards of political reasonableness.
Carl Larson, PhD, is professor emeritus of human communication studies at the University of Denver. His research focuses on groups and teams and on negotiation and collaborative processes. He has co-authored seven books, including *Teamwork: What Must Go Right/What Can Go Wrong* (with Frank LaFasto), *Collaborative Leadership* (with David Chrislip), and *When Teams Work Best* (with Frank LaFasto).

Christopher Nelson, PhD, teaches statistics at the Daniels College of Business and research methods at the Morgridge College of Education at the University of Denver. He is also a lead consultant at Nelson Consulting, LLC. Recent publications appear in *Journal of Colorado Libraries* and the *Proceedings of The International Communication Association’s 2007 Convention*.

David L. Olds, PhD, is professor of pediatrics, psychiatry, preventive medicine, and nursing at the University of Colorado Health Sciences Center, where he directs the Prevention Research Center for Family and Child Health. He has devoted his career to investigating methods of preventing health and developmental problems in children and parents from low-income families. He has published widely in pediatrics, prevention science, and community psychology.

Erik Johnston, PhD, is an assistant professor of policy informatics in the School of Public Affairs at Arizona State University. He studies how models and simulations can aid individuals and groups make and evaluate informed policy judgments. He is interested in the implementation process of civic collaboration, alternative methods of health care delivery, and collaborative decision making.