

*This section includes shorter papers (e.g., 10-15 double-spaced manuscript pages or less) describing methods and techniques that can improve evaluation practice. Method notes may include reports of new evaluation tools, products, and/or services that are useful for practicing evaluators. Alternatively, they may describe new uses of existing tools. Also appropriate for this section are user-friendly guidelines for the proper use of conventional tools and methods, particularly for those that are commonly misused in practice.*

## Measuring Collaboration Among Grant Partners

Bruce B. Frey  
Jill H. Lohmeier  
Stephen W. Lee  
Nona Tollefson  
*University of Kansas*

**Abstract:** Collaboration is a prerequisite for the sustainability of interagency programs, particularly those programs initially created with the support of time-limited grant-funding sources. From the perspective of evaluators, however, assessing collaboration among grant partners is often difficult. It is also challenging to present collaboration data to stakeholders in a way that is meaningful. In this article, the authors introduce the Levels of Collaboration Scale, which was developed from existing models and instruments. The authors extend prior work on measuring collaboration by exploring the reliability of the scale and developing a format for visually displaying the results obtained from using the scale.

**Keywords:** *collaboration; grant partners; cooperation; visual displays*

Evaluators are routinely called on to assess interactions among different groups and agencies involved in large-scale grant program activities. An increased level of *collaboration* is frequently an explicitly identified objective (and almost always an implicit grant objective) and, therefore, a target to be measured by program evaluators. Collaboration among the various organizations and individuals involved in these multifaceted approaches is

---

Bruce Frey, PhD, University of Kansas, Psychology and Research in Education, 1122 West Campus Road, Room 643, Lawrence, KS 66045; phone: (785) 864-9706; fax: (785) 864-3820; e-mail: bfrey@ku.edu.

Authors' Note: This article was prepared under a Safe Schools, Healthy Students grant from the U.S. Department of Education, the U.S. Department of Health and Human Services, and the U.S. Department of Justice.

viewed as essential for success (Center for Mental Health in Schools, 2003; Gajda, 2004; Grubbs, 2000; Riggins, 2004). Typically, it is hoped that collaborations formed for the purposes of implementing grant-funded programs will remain after the funding ceases. Collaboration is seen as a prerequisite for the sustainability of interagency programs (Hogue, 1993; Perkins, 2002; Peterson, 1991), particularly for those programs initially created with the support of time-limited funding sources.

From the perspective of evaluators, however, assessing collaboration is often difficult. Models of collaboration among agencies, groups, and community stakeholders are notoriously difficult to translate into valid and reliable instruments that can measure meaningful change in the level and pattern of collaboration. This article reviews the existing stage models of collaboration, presents an instrument for assessing collaboration that is an extension of those models, examines preliminary reliability and theoretical validity evidence for the instrument, and suggests the use of a method of graphical display to map the sometimes elusive nature of collaboration. The use of these tools in a local Safe Schools, Healthy Students evaluation is discussed.

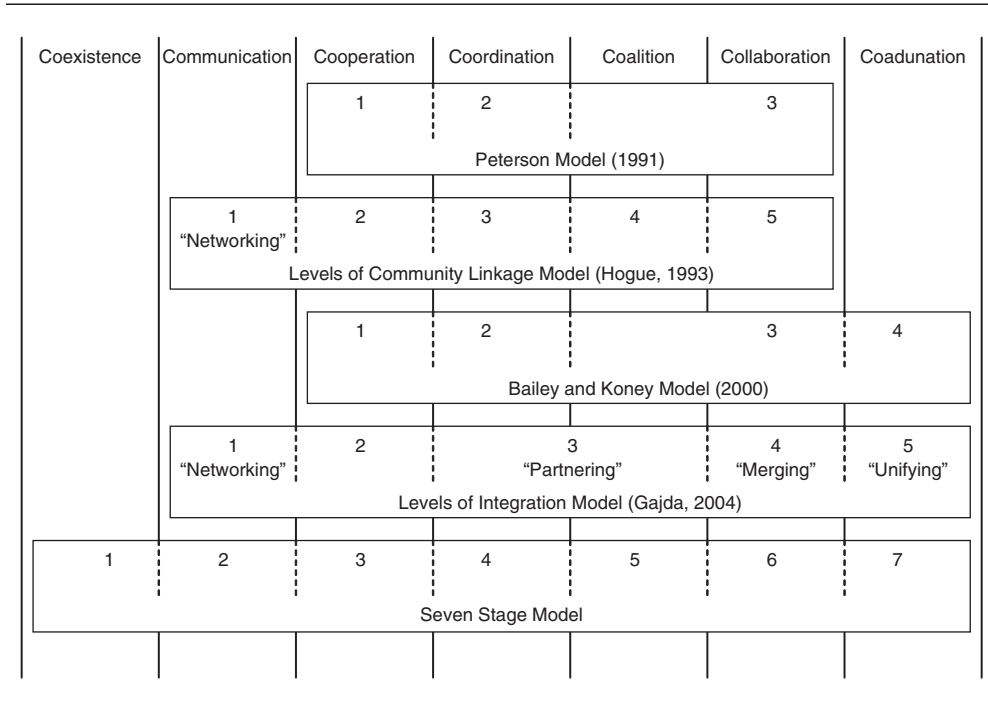
### Stages of Collaboration

Collaboration has a variety of definitions and names but is generally treated as meaning the cooperative way that two or more entities work together toward a shared goal. Collaboration among individuals with shared goals in professions such as mental health and education has been studied (Kabler & Genshaft, 1983; Moriarty, 2000; Smith, Frey, & Tollefson, 2003), as has collaboration within and among individuals in the development of small groups (Tuckman, 1965; Tuckman & Jensen, 1977). Collaborative interactions among businesses as a profit strategy and models of the dynamics of those relationships have been explored in the business management and networking literature (e.g. Buckley & Casson, 1988; Todeva & Knoke, 2005). Additionally, some researchers have explored the specific nature of successful relationships within school and business partnerships (Ash, 1989; DelPizzo, 1990; Kysiak, 1986; Rockefeller, 1986). Collaboration among public and nonprofit organizations, the alliance most typically seen by evaluators, has become a key issue in public administration research (Grubbs, 2000). However, a comprehensive theory of collaboration within the types of shared organizational efforts (or *strategic alliances*; Gajda, 2004; Todeva & Knoke, 2005) formed through grant-funded initiatives and other public service efforts has not been presented in the literature.

Todeva and Knoke (2005) argued that the principal dimensions that classify levels of collaboration among organizations (in their research, businesses) are increasing integration and increasing formalization. A similar description of moving from loose integration to greater consolidation in collaborative alliances between government and nonprofits was provided by Arsenault (1998). Bardach (1999), also interested in government and nonprofit interactions, suggested that the level of integration can be quantified using a concept of *collaborative capacity* that takes into account the level of integration and social costs.

Preliminary models of collaboration within social-service-oriented alliances have been presented (Bailey & Koney, 2000; Gajda, 2004; Hogue, 1993; Peterson, 1991). These models commonly focus on *stages* of collaboration through which interagency initiatives might move. Gajda (2004) argued that groups will pass from lower to higher stages of collaboration before they can be effective. These stage theories describe levels of collaboration, with the lowest level being little or no collaboration and the highest level being full collaboration or, ultimately, complete unification. The models differ on the number of stages, the range of levels included, and the definitions of various stages, but they have much in common.

**Figure 1**  
**Stage Models of Collaboration**



Peterson (1991) proposed three types of agency interaction: cooperation, coordination, and collaboration. Although described by Peterson as distinct states of interactions among agencies and not offered as a strict series of stages, in Gajda’s (2004) review of Peterson’s model, they are presented as a three-point continuum. These categories are differentiated on the basis of the degree of member autonomy associated with each. Hogue (1993) suggested five levels of *community linkage*: networking, cooperation or alliance, coordination or partnership, coalition, and collaboration. The levels differ by purpose, the structure of decision making, and the nature of leadership. Bailey and Koney (2000) offered a model similar to these, with four steps ending with complete unification: cooperation, coordination, collaboration, and coadunation (which means “having grown together”). A five-stage model consistent with previous stage approaches was suggested by Gajda. The *level-of-integration* model has five ordered steps: networking, cooperating, partnering, merging, and unifying. The steps differ on purpose, tasks and organizational strategies, leadership and decision making, and type and frequency of communication.

A summary and comparison of the various stage approaches to collaboration among groups offered in the literature is shown in Figure 1. Uniform terms are used to label stages, with terminology specific to each model indicated where necessary. Figure 1 includes a seven-stage model, which simply extends the previously identified stages to include the possibility that while both groups exist, there may be no collaboration whatsoever between them. Those interested in assessing collaboration might best include this lowest possible score or category of collaboration, which represents the state of participating agencies before any level of collaboration has begun.

## Levels of Collaboration Survey

Safe Schools, Healthy Students grants, first funded in 1999 through the U.S. Department of Education, the U.S. Department of Justice, and the U.S. Department of Health and Human Services, have been awarded to hundreds of cities and school districts over the past 5 years and include evaluation mandates. Although each local grant varies, ours was typical in that it covered objectives in a range of health and safety domains, such as mental health, drug and alcohol use, violence, and school readiness. The evaluation of a Midwest school district's implementation of a Safe Schools, Healthy Students initiative included the task of measuring levels of collaboration among grant partners. The Safe School, Healthy Students requests for proposals viewed sustainability as a necessary ingredient for successful initiatives. Consequently, the grant partners subsequently identified increased collaboration as a key goal for the initiative.

Our approach to this measurement task was to first define the construct of collaboration, through a search of the published literature and through interviews with representatives of various participating grant partners. The literature review is summarized earlier in this article.

Grant partners—heads or representatives of the various agencies and community groups receiving funding under the grant—were solicited for interviews about their perceptions of the meanings of several key terms and abstract outcomes for the grant. These terms included *safe schools*, *healthy students*, and *collaboration*. Participant interviews revealed that collaboration was perceived as a level of cooperation that involves teamwork, communication, and consideration. The working definition jointly developed by the evaluation team and the grant partners was that collaboration was a variety of parties coming together to reach a shared goal.

Next, we identified a model to provide the conceptual framework for any instrumentation we chose or developed. The five stages of Hogue's (1993) model of levels of community linkage were chosen as the most relevant framework. In addition to Hogue's levels, it was possible that some partner groups had no interaction with other groups, especially at baseline, and this possibility was reflected in the final instrumentation by allowing respondents to choose "0" to indicate no collaboration whatsoever.

In developing the survey instrument, the detailed descriptions of the levels of community linkage provided by Hogue (1993) and discussed by Borden and Perkins (1998, 1999) were combined and shortened. Whereas combining the subcomponents of linkages provided in Hogue's full framework has the potential for losing interesting information, concentrating on the common threads that define the stages allowed us to provide simpler definitions for respondents. Although Hogue's full model treats collaboration as being represented in three different dimensions (purpose, structure, and process) at each of the five levels, for the sake of parsimony, our instrument asked about the six levels only, without further breakdown into the three different dimensions of collaboration.

Given the definitions of each level, respondents were asked to what extent they collaborated with each other grant partner. Answer options were on a scale ranging from 0 to 5, with 0 indicating *no interaction at all* and 5 indicating the *collaboration* level, using Hogue's (1993) taxonomy. The instrument, with descriptions of each stage, is shown in Figure 2. For the purposes of this article, the name of each group or organization has been replaced with a generic descriptive name.

Using the Levels of Collaboration Scale as part of our evaluation, data were collected during the baseline year, before regular monthly meetings with a stable group of representatives from each of the partner organizations that had been established, and respondents were asked to fill out a survey at the end of one-on-one interviews that were conducted as part of the start-up evaluation activities. Respondents were usually the heads of the agencies or projects, who

**Figure 2**  
**Levels of Collaboration Survey**

This form is designed for those who work in one of the organizations or programs that are partners in the *Safe Schools, Healthy Students* initiative. Please review these descriptions of different levels of collaboration.

- On the response section at the bottom of the page, please circle the name of the organization or group with which you are associated.
- Using the scale provided, please indicate the extent to which you currently interact with each other partner. (Skip your own row.)

Relationship Characteristics	Five Levels of Collaboration and Their Characteristics				
	Networking 1	Cooperation 2	Coordination 3	Coalition 4	Collaboration 5
	-Aware of organization roles -Loosely defined roles -Little communication -All decisions are made independently	-Provide information to each other -Somewhat defined roles -Formal communication -All decisions are made independently	-Share information and resources -Defined roles -Frequent communication -Some shared decision making	-Share ideas -Share resources -Frequent and prioritized communication -All members have a vote in decision making	-Members belong to one system -Frequent communication is characterized by mutual trust -Consensus is reached on all decisions
<i>Safe Schools, Healthy Students</i> Partners	No Interaction at All	Networking	Cooperation	Coordination	Coalition
Mental Health Agency	0	1	2	3	4
Early Childhood Programs	0	1	2	3	4
Parent Education Program	0	1	2	3	4
School District Prevention Counselors	0	1	2	3	4
After School Programs Director	0	1	2	3	4
Student Improvement Teams	0	1	2	3	4
Principals	0	1	2	3	4
Teachers	0	1	2	3	4
Police Department	0	1	2	3	4

would later attend the monthly grant partner meetings. Those available to generate test-retest reliability data were visited a second time about a month later. At the end of the first year, the instrument was administered as part of the monthly grant partner meetings and again at the next meeting to gather test-retest reliability data.

## Collaboration Map

Data collected with the Levels of Collaboration Scale can be reported quantitatively using different formats and different summations depending on the interests of evaluators, grant directors, and stakeholders. Collaboration can be reported as the mean level of perceived collaboration across all respondents for all partners, summarized in other meaningful ways, or provided as raw data in a table. Because by definition, collaboration exists only when two or more parties interact with each other, situations in which two partners report different levels of collaboration with each other represent areas for exploration and discussion between those partners. Because the scale assesses perceptions of collaboration, different perceptions by two collaborating partners may both be valid responses.

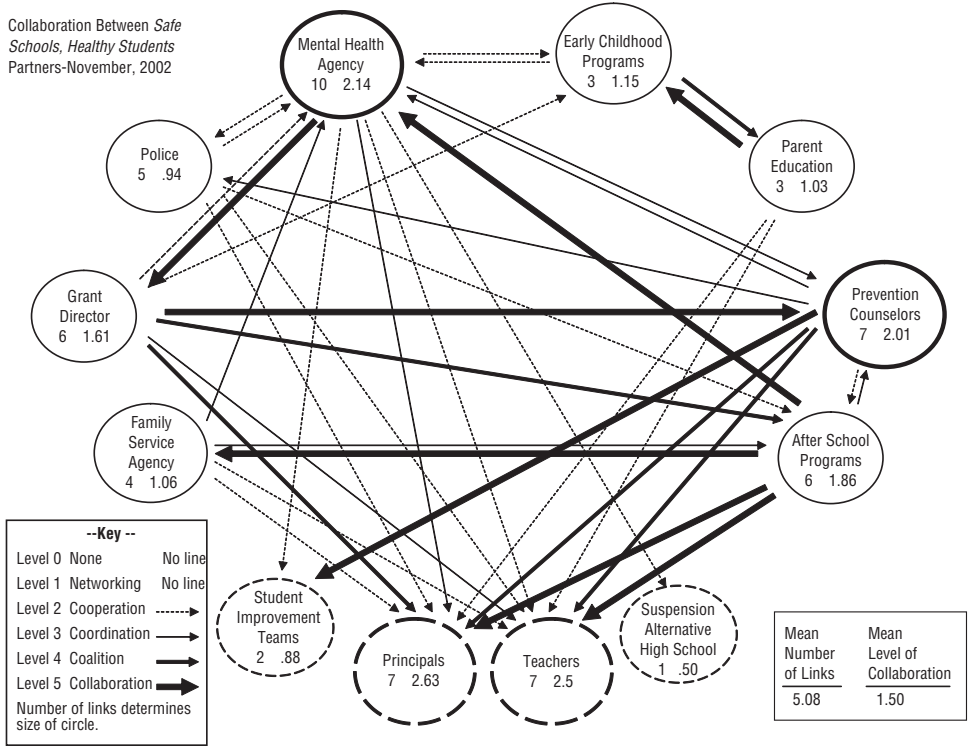
Complex information such as the data generated by our scale is difficult to communicate but can be made clear with a good picture. As the amount of complex data available to evaluators has increased, the creative use of data displays for quantitative information has become essential to promote the ease of interpretation (e.g. Tufte, 1996, Tukey, 1977, 1989; Wainer, 2005). Because patterns of collaboration among multiple entities may be complex, data on collaboration may best be represented visually. The idea of displaying levels of collaboration between entities using circles to represent partners and lines between the circles to represent collaboration was included in a Safe Schools, Healthy Students conference paper by Cross (2003). We adapted and extended that method, added new elements, and produced collaboration maps from the data produced with the Levels of Collaboration Scale. A sample map is shown in Figure 3.

In our collaboration maps, connecting arrows between circles (partners) indicate that partners reported moderate or high levels of collaboration with one another. On the map, collaboration levels reported as 0 or 1 are not represented with lines. This allows for four types (or thicknesses) of lines to be used to represent the variety of levels, with thicker lines indicating higher levels of collaboration. The two numbers in each circle reflect, first, the number of partners with whom each entity collaborates and, second, the mean level of collaboration across all partners. The sizes of the circles are based on the levels of collaboration with other groups, as represented by the number of links with other circles. Size criteria and the number of sizes used depend on the variability across partners. For these data, we use three relative circle sizes: small for one or two links, medium for three to six links, and large for seven links or more. We also use broken or dotted lines for four circles to indicate groups that did not respond to the scale but were response options. In other words, arrows could go toward them but not come from them.

We chose to use two lines to indicate each collaborative relationship, because respondents at both ends of the dyads may have different perceptions of their collaborative levels. Another option would be to use only one line with two arrows and define the value for that line (its thickness) using the mean of both respondents' scores. However, using this approach would reduce the interpretative value of the nature of the collaboration between the entities.

The map includes additional informative elements. In each circle is the mean level of collaboration reported going to and from each grant partner, as well as the number of links to that circle. This provides individualized collaboration information. Those individual group data can also be averaged across all partners to produce the mean level of collaboration among all partners, as well as the mean number of links among partners. In this way, different

**Figure 3**  
**Collaboration Map**



perspectives on how to reflect collaboration are provided. Is it the mean overall level that is the key or the number of connected partners?

Although the example shown is complex, after one is familiar with the symbols, areas of high and low collaboration become apparent, as do areas of disagreement. The basic format can be adapted easily to emphasize additional or different information. For example, with plenty of space and fewer partners, one can arrange the circles so that the distance between them represents levels of collaboration as well. Whatever the method, collaboration maps allow for interpretations of collaboration from a variety of perspectives. Change in the number of connecting lines, their thicknesses, and even the number of circles can represent changes in collaboration across time; this method also allows for a visual representation of *systems change*, an outcome that in our experience is commonly evaluated in school reform initiatives.

### Discussion

The use of the Levels of Collaboration Scale is consistent with all the major stage theories of interagency collaboration and identifies the levels of collaboration that are common across them. Allowing for a full range of scores to represent all possible levels of interaction provides some theoretical evidence of validity for the measurement approach. Interestingly, in our experience, the presence of the full range of scores on the survey has led to concerned discussions



**Table 1**  
**Preliminary Test-Retest Reliability Estimates (1 month)**

Data Collection Occasion	Number of Participants	Test-Retest Correlation <i>M</i> ( <i>SD</i> )
Start of Year 1	2	.85 (.05)
End of Year 1	9	.87 (.09)
End of Year 2	8	.81 (.29)
End of Year 3	9	.85 (.08)

about the “correct” hoped-for level of collaboration. Most cooperating groups do not wish to merge or unify or become one, and the presence of such high levels on the scale suggests to some that they are expected to want to attain that level of collaboration. We make it a point now to explain to grant partners that their legitimate goals for collaboration might well be much more moderate, and in our experience, they often are.

The reliability of the scores produced by the Levels of Collaboration Scale was a concern of ours in our local evaluation, and we collected some data to address it. Because we were measuring change, our primary reliability concern was test-retest reliability, or the stability of scores across brief periods of time. A high consistency in responses in the absence of real change is essential for the method to be sensitive enough to detect real change when it occurs. As one might expect, the number of participating agencies in a multiagency grant collaboration will be small. Consequently, the stability of collaboration scores is particularly crucial. To assess test-retest reliability, we conducted small psychometric studies during each of the first 3 years of the administration of the scale in our local evaluation. With only about 10 grant partners participating in the collaboration data collection, our sample sizes were quite small. A summary of the test-retest reliability data grant partners produced is presented in Table 1.

During the baseline year, 2 key respondents were asked to respond to the form twice, with an intervening time period of 1 month. On the second administration, they were asked to refer to the originally referenced time point to control for any actual collaborative changes that may have occurred during the early weeks of preliminary meetings and planning. Test-retest reliability for the small sample was high. Correlations of stability for the 2 respondents, each providing nine pairs of matched scores, were .80 and .90, for a mean value of .85. When collaboration data were collected at the start of the 2nd year, test-retest data were collected from all key respondents. Correlations of stability for respondents ranged from .69 to .97, with a mean reliability coefficient of .87 ( $SD = .09$ .) Test-retest reliability data were again collected during the third annual administration of the survey. Data from 8 respondents produced a mean test-retest correlation of .81 ( $SD = .29$ ). On a fourth occasion, 1 year later, reliability estimates were again encouraging, with a mean  $r$  value of .85 ( $SD = .08$ ,  $n = 9$ ). Although each reliability study was small, the consistency of results above .80 suggests that the instrument can produce stable scores. Larger scale studies are necessary before stronger conclusions can be made, however.

The Levels of Collaboration Survey provides definitions of stages of collaboration to respondents. The definitions are consistent with the scholarly literature and theory that support the assumption of general construct-based validity for the scale. Scores on the scale are meant to represent the amount of collaboration between organizations at a given time, and scores are intended to be used to reflect change over time. Two specific validity concerns relate to those stated purposes of the assessment.

First, when participants respond to the scale, are they responding on behalf of the entire organizations they represent or reflecting only their personal, individual experiences? This is an important question conceptually because the core of collaborative relationships among



groups is the collaborative relationships between individuals who are part of those groups. Gajda (2004) pointed out that collaboration depends on “positive personal relations and effective emotional connections between partners” (p. 69). In our early use of the instrument, we spoke with respondents about the referents they chose when responding. About half assumed that they should report as individuals; the rest reported as representatives of their groups or used some combination of roles. After the first use of the survey, we recommend that respondents be sampled from organizations and then be instructed to respond as individuals. In some cases, though, an individual is, essentially, an entire organization, at least in terms of being the only one who interacts with other grant partners. In those cases, the issue of whether respondents are individuals or representatives of groups is of only abstract, theoretical interest.

A second concern is whether the instrument can detect change. Evaluators are constantly faced with the expectation that important changes in outcomes will be detected using a given measurement system. The preliminary evidence of high test-retest reliability of the Levels of Collaboration Scale indicates a good degree of precision in measurement, suggesting that it is an appropriate tool to measure change. When the scale was used in our local evaluation, with just seven representatives on both measurement occasions, a change in mean collaboration of 0.55 standard deviations was observed between baseline and the end of the 1st year of grant activities. The mean moved from 1.40 ( $SD = 0.55$ ) to 1.71 ( $SD = 0.57$ ) during the 1st full year of the grant. Taking into account all respondents, not just those responding on both occasions, the mean level of collaboration moved from 1.50 ( $SD = 0.54$ ) to 1.77 ( $SD = 0.50$ ). Before the grant initiative began, partners were above the networking level of collaboration and after a year had moved toward the cooperation level. Early in our meetings with grant partners, we urged them to set a collaboration goal, using the scale of 0 to 5. Establishing a quantifiable target allows for the reporting of progress toward that target. Grant partners had chosen a goal of 2.41 on the scale (somewhere between *cooperation* and *coordination*), and the data generated by the survey can be interpreted as indicating that they were 30% toward their goal. The scale appears to provide meaningful information about change in reported levels of collaboration over time.

Without independent studies of the nature of the construct being assessed, however, the validity of the measurement system cannot be fully established. Future exploration of the validity of the instrument should include both qualitative studies to examine the nature of the construct from the perspective of respondents and correlational studies examining the relationships between scores on the Levels of Collaboration Scale and scores on other collaboration measures and criteria.

Future study of the scale, survey, and map is needed, along with refinement, especially in the areas of how the use of these reporting tools affects grant partner behaviors and decisions. Because there is a likelihood that the high-end anchor, involving the complete merging of the groups, is unrealistic for most grant partners, it might pay to widen the continuum of possible points in the middle of the scale to allow for more precision in that crucial range.

Our experience with the survey and maps in our local evaluation has been that stakeholders, respondents, district administrators, teachers, principals, and grant partners find the information useful and persuasive. Even more, the visual representation method, used as feedback, has resulted in grant partners actually identifying collaboration goals and targets that were not part of the original design. In this context, the scale operates as a formative assessment. In addition, the level of collaboration might well have increased to a small degree because of the interactions and discussions about Levels of Collaboration scores themselves. As with many assessment activities in program evaluation, the data collection itself has led to system changes. The large amount of information presented by the graphic technique makes for a powerful reporting tool (and, in the case of our local evaluation, could even act as an intervention) and

provides a convenient visualization of the dynamics of collaboration. Furthermore, including data collection on the level of collaboration between grant members can add to the developing theory and research in systems-level collaboration.

## References

- Arsenault, J. (1998). *Forging nonprofit alliances*. San Francisco: Jossey-Bass.
- Ash, A. M. (1989). *Interorganizational relations and effectiveness in school-business partnerships*. Unpublished doctoral dissertation, University of Miami.
- Bailey, D., & Koney, K. (2000). *Strategic alliances among health and human services organizations: From affiliations to consolidations* (Abr. ed.). Thousand Oaks, CA: Sage.
- Bardach, E. (1999). *Getting agencies to work together: The practice and theory of managerial craftsmanship*. Washington, DC: Brookings Institution.
- Borden, L., & Perkins, D. (1998). Evaluating your collaborative effort. *Program Evaluation Newsletter*, 1, 5.
- Borden, L., & Perkins, D. (1999). Assessing your collaboration: A self-evaluation tool. *Journal of Extension*, 37(2), 67-72.
- Buckley, P., & Casson, M. (1988). A theory of cooperation in international business. *Management International Review, Special Issue*, 19-38.
- Center for Mental Health in Schools. (2003). Safe students, healthy schools: A collaborative process. *Addressing Barriers to Learning*, 8(2), 1-7.
- Cross, J. E. (2003, April). *Evaluation instruments that measure change, growth and collaboration*. Paper presented at Safe Schools, Healthy Students, Arlington, VA.
- DePizzo, M. T. (1990). *A naturalistic study of the salient themes of a school/business partnership*. Unpublished doctoral dissertation, West Virginia University.
- Gajda, R. (2004). Utilizing collaboration theory to evaluate strategic alliances. *American Journal of Evaluation*, 25(1), 65-77.
- Grubbs, J. W. (2000). Can agencies work together? Collaboration in public and nonprofit organizations. *Public Administration Review*, 60(3), 275-280.
- Hogue, T. (1993). *Community-based collaboration: Community wellness multiplied*. Bend, OR: Chandler Center for Community Leadership. Retrieved April 21, 2004, from <http://crs.uvm.edu/ncco/collab/wellness.html>
- Kabler, M., & Genshaft, J. (1983). Structuring decision-making in multidisciplinary teams. *School Psychology Review*, 12(2), 150-159.
- Kysiak, R. C. (1986). Role of the university in public-private partnerships. *Proceedings of the Academy of Political Science*, 36(2), 47-59.
- Moriarty, M. L. (2000). *Attitudes on collaboration in education: A pilot study*. Unpublished manuscript, University of Kansas.
- Perkins, T. (2002). *Comprehensive community initiatives (CCI): A comparison of community implementation plans*. Lincoln: University of Nebraska Public Policy Center. Retrieved October 1, 2003, from [http://ppc.unl.edu/reports\\_publications/csipcommunity\\_comparison.pdf](http://ppc.unl.edu/reports_publications/csipcommunity_comparison.pdf)
- Peterson, N. L. (1991). Interagency collaboration under Part H: The key to comprehensive, multidisciplinary, coordinated infant/toddler intervention services. *Journal of Early Intervention*, 15(1), 89-105.
- Riggins, J. M. (2004, April). *Making a mess of program theory: Evaluating complex, collaborative initiatives*. Paper presented at the annual meeting of the American Educational Research Association, San Diego, CA.
- Rockefeller, D. (1986). Ingredients of successful partnerships: New York City case. *Proceedings of the Academy of Political Science*, 36(2), 122-154.
- Smith, S. J., Frey, B. B., & Tollefson, N. (2003). A collaborative cohort approach to teacher education: Modeling inclusive practices. *Action in Teacher Education*, 24(1), 55-61.
- Todeva, E., & Knoke, D. (2005). Strategic alliances and models of collaboration. *Management Decision*, 43(1), 123-148.
- Tuckman, B. W. (1965). Developmental sequence in small groups. *Psychological Bulletin*, 63(6), 384-399.
- Tuckman, B. W., & Jensen, M. A. C. (1977). Stages of small group development revisited. *Group and Organizational Studies*, 2(4), 419-427.
- Tufte, E. R. (1996). *Visual explanations*. Cheshire, CT: Graphics Press.
- Tukey, J. W. (1977). *Exploratory data analysis*. Reading, MA: Addison-Wesley.
- Tukey, J. W. (1989). Data-based graphics: Visual display in the decades to come. *Proceedings of the American Statistical Association*, 84, 366-381.
- Wainer, H. (2005). *Graphic discovery*. Princeton, NJ: Princeton University Press.