

## **6. Evaluation in All its Complexity**

In this chapter we will, on the basis of the discussion thus far, propose a new framework for the evaluation of community-based HIV prevention. The first step will be to present a series of propositions which describe the necessary characteristics of such an approach, reflecting the issues raised by the sexual transmission of HIV (Chapter 2), the practice of community-based HIV prevention (Chapter 2), and current problems in evaluation practice (Chapter 3). We will then describe how complexity theory can be used as a meta-theory to fulfill these propositions (Chapters 3 and 4). Finally, we will propose the primary features required of a complex theory of evaluation research.

### **6.1 Basic Characteristics Required of Evaluation Theory and Practice**

In order to respond adequately to the phenomenon of HIV transmission and the community structures designed for its prevention, the theory and practice of evaluation research needs to take into account the following aspects:

#### **6.1.1 A Multiplicity of Causal Factors at Various Levels**

Evaluation research needs to recognize that HIV epidemics are caused by a host of factors at various levels in society, thus being open in principle to interventions of various kinds. An exclusive focus on programs designed to affect the behavior of individuals unnecessarily narrows the scope of activities which can be evaluated. Particularly, research examining the structural level of intervention needs to be strengthened, in spite of the considerable methodological barriers which this presents.

#### **6.1.2 The Unique Role of Community-Based Organizations**

The unique role of community-based organizations in HIV prevention needs to be recognized. Above all, this means acknowledging the centrality of social change within the goals of community-based prevention activities. Thus, interventions need to be examined within the context of such goals and not reduced to a handful of factors measurable at the individual level.

#### **6.1.3 The Interests of Stakeholders**

To be truly responsive to the needs of community-based groups currently active in HIV prevention practice, evaluation research should be guided by the concrete

and pressing issues impinging on stakeholders. Evaluation is not a goal unto itself, but rather a tool for improving prevention theory and practice by providing a systematic examination of services. This means assisting the various stakeholders in formulating their questions and in developing appropriate study protocols which address these questions.

#### 6.1.4 Formative, Process, and Outcome Evaluations

Evaluation does not begin at the completion of an intervention, but needs to accompany all stages of program development. It cannot be concluded that a program has “worked” if sufficient information about the spread of HIV among the target population as well as information about program delivery are not available. Evaluation thus has the function of unifying often disparate activities within an organization in the interest of systematically examining the effects of prevention work.

#### 6.1.5 A Selection of Methods Appropriate to Setting

Given the multi-dimensional nature of the problem of HIV transmission, the multi-level nature of intervention, and the diversity of practice settings, interventions, and target populations, evaluation research cannot privilege particular research methods over others when gathering and interpreting data. The choice of method should be guided by the unique characteristics of the program to be evaluated, the questions of the stakeholders, and the constraints imposed by setting and practice structures. Ideally, the researcher will have recourse to various sources of information about the program of interest and be able to employ a variety of data-gathering methods so that he or she can triangulate findings.

#### 6.1.6 A Selection of Measures Appropriate to Goals and Structures

Likewise, the selection of measures needs to reflect the goals and structures of the prevention work as well as the questions of interest to the stakeholders. Standardized measures and other recognized research tools can be employed to the extent that their validity in the situation at hand can be established. This means more than validity in a strict scientific sense; the instruments used need to make sense to the organization being evaluated and reflect their understanding of the work. Difficulties in defining and operationalizing important variables should not lead the researcher to reduce the evaluation to a more narrow scope, but rather to the development of new

tools in order to describe hard-to-measure constructs. Comprehensiveness in terms of addressing an organization's prevention goals should take precedent over methodological ease.

#### 6.1.7 Evaluation as a Tool in Community Development

The highest priority in the evaluation of community-based organizations needs to be community development. This is reflected in the focus on stakeholder interests and the "service oriented" approach mentioned in Chapter 3. In addition, a successful collaboration between researchers and community groups needs to be established. The principles defined by the International Network for Community-Based Research on HIV/AIDS offer a basis for such a working relationship. At the very least, such a collaboration should result in evaluation reports which are readily understandable to all parties involved and which include clear, realistic recommendations for the improvement of prevention practice. Ideally, collaboration between researchers and community organizations will be part of a larger organizational change process with the goal of improving prevention services.

#### 6.1.8 Evaluation Research, Best Practice, and Scientific Debate

As a scientific pursuit, evaluation research is obligated to work toward the accumulation of knowledge in the field, thus going beyond the immediate interests of stakeholders. This means publishing findings and keeping abreast on developments in the field as a whole. Although adapted to specific local needs, evaluation reports contain information which can be useful to a broader audience. An important task of the researcher is to identify this information and to make it known.

#### 6.1.9 The Unique Aspects of the German Discourse

The desire to participate in the international discussion on evaluation and HIV should not overshadow the importance of German traditions in social reform and evaluation. The published literature is largely from North America and thus needs to be examined critically to determine what is applicable to the discourse here. Most importantly, evaluation research of German AIDS service organizations needs to be integrated into the political and practical issues of most relevance to the development of prevention in this country.

## 6.2 Complexity Theory as a Unifying Meta-Theory

At this point, we could simply make concrete recommendations for the improvement of evaluation research based on the above-named principles. This would, however, not bring us much further in the debate surrounding evaluation practice as described in Chapter 3. The largely practical considerations summarized above do not offer a clear alternative to those who argue for experiments and evidence in the sense of the dominant scientific paradigm. Promoting the triangulation of methods, a “service-oriented” evaluation style, and the recognition of multiple causes does not address the underlying epistemological doubt concerning non-experimental data.

Complexity theory provides a basis for conceptualizing HIV prevention which extends beyond “the conventional view” but without rejecting either rationalism or empiricism as a basis for future work. Complexity theory can be viewed as a meta-theory which unifies the various observations presented in the preceding chapters. Pitting quantitative against qualitative methods or surveillance data against needs-based analyses or experimental methods against other means of data analysis represents a destructive and artificial polarization of perspectives. As described in Chapter 4, complexity theory *requires* a plurality of methods and points-of-view for the very reason that the subject of research—complex human systems—exceeds the capacity of any one approach or discipline. Complexity theory can best be viewed as a framework in which all disciplines can continue their work and at the same time—and most importantly—communicate with other disciplines in order to produce comprehensive explanations of human problems and interventions to address them. In this vein, the following summary of the primary characteristics of complexity theory represents the backbone of this unifying framework.

### 6.2.1 Complexity as a Norm in Human Systems

Complexity is a norm in human systems. When thinking in terms of complexity theory, all human systems are viewed as being complex in nature. This results in the scientific imperative to look for the various features of complexity when conducting research on human problems. Thus, the researcher is not content with one level of explanation, but seeks to identify the hierarchy of levels and their interconnectedness. Even when focusing on a particular aspect of a problem, the researcher remains aware that this is only one piece in a larger puzzle. And, from a complexity perspective, it is

the larger puzzle and the underlying dynamic which is of primary interest. As was discussed in Chapter 2, the complex nature of HIV transmission has become apparent, in spite of a science which has tended to isolate certain aspects from others. Had a complexity approach been applied from the beginning, such an isolation would not have been possible.

### 6.2.2 The Common Features of Complex Systems

Complex systems have common features which can guide theory and practice. Not only are human systems by their very nature complex, they have features which one finds in all complex systems, as described in Chapter 4. When applying complexity theory to HIV prevention we are prompted to look for these various features in order to understand a system of HIV transmission as a whole.

In applying complexity theory to the spread of HIV within the male prostitution scene in Germany, we used the concept of a nested hierarchy of causal levels to bring more clarity to the structural understanding of causes already in use. Further, we constructed a dynamic model, thereby directing our focus to changes over time. This led us to look for signs of the overall transmission dynamic within the system, thus going beyond an analysis of the contribution of individual parts. Finally, the focus on the dynamic provided ideas for how prevention could be organized to have the greatest effect on prevalence in the long run.

By providing a framework for conceptualizing complex systems we are better able to generate and integrate knowledge about the systems as a whole. Claims of connections between macro and micro levels take on a clearer structure and meaning.

### 6.2.3 Multiple Sources of Knowledge

Knowledge from various sources is needed in order to understand complex systems. When conceptualized in a complex way, the need to draw information from various sources regarding any given system of HIV transmission becomes readily apparent. The debate regarding randomized clinical trials can be put into perspective as we realize that, even in the best of circumstances, experiments alone do not clarify how best to intervene in the larger dynamic. Thus qualitative and quantitative, experimental and non-experimental, intervention-orientated and basic research are all needed to put together a picture of how transmission occurs in a particular place and time and what options may provide the most impact. This need for various types of

input also makes clear the value of the various stakeholders' perspectives in generating information. Thus, triangulation is seen not as an alternative method, but as an epistemological imperative.

#### 6.2.4 Dynamic Models

Dynamic models are indispensable. The agent-based simulation exercise in Chapter 5 makes clear the importance of generating a dynamic model of the problem at hand. The use of computer-based simulations may make certain aspects of the underlying dynamic more readily apparent and thus accessible to analysis. As this new modality continues to develop, its application will become more practical. In any case, it is important to construct a model of the problem which is not static but rather, to paraphrase Dietrich Dörner, has the character of a motion picture. The element of time is central when considering how the various levels of the model interact to produce the outcome of interest.

#### 6.2.5 Appropriate Strategies

Complex systems can be influenced when approached with an appropriate strategy. Dietrich Dörner provides an insightful and comprehensive summary of how best to approach a complex system strategically (Chapter 4). Through his work we see how complexity theory is not only a basis for conceptualizing complex problems, but also for developing appropriately complex ways of responding to them.

### ***6.3 Primary Features of a Complexity-Based Evaluation Theory and Practice***

By integrating the first two sections of this chapter we can outline the primary features of a complexity-based theory and practice for the evaluation of community-based HIV prevention. This will be organized under the following headings: Defining the focus of research; formative evaluation; process evaluation; and outcome evaluation.

#### 6.3.1 Defining the Focus of Research Through Collaboration

The basis for an effective evaluation is a close collaboration between the researcher and the community-based organization. The principles defined by the International Network for Community-Based Research on HIV/AIDS offer a basis for es-

tablishing such a working relationship. The goal of the collaboration is to produce scientifically sound information which will be useful for the improvement of prevention services. In negotiating this relationship, the following questions are central:

- Who are the stakeholders and what are their interests in the evaluation process?
- What questions are of interest to the stakeholders?
- What resources are available for evaluation activities?
- What structures are currently employed in prevention work?
- What are the overall prevention goals of the organization? Is there an underlying general theory or approach?

Through a careful consideration of these topics, the researcher and community-based organization can establish a working agreement which specifies the questions to be answered and the scope of the evaluation project. The answers to the above questions are often not readily apparent: In many cases, they have not been clearly articulated by the organization or they constitute a consensus among staff which has never been formalized. By initiating a discussion about these questions, the researcher is intervening in the organizational system in the interest of clarifying what prevention means to staff members in the overall context of the organization's mission.

### 6.3.2 Formative Evaluation

In the strict sense, a formative evaluation should be conducted before any prevention work is done. It is the formative evaluation which provides the necessary information about the transmission of HIV in order to plan a potentially effective prevention strategy. In practice, many community-based organizations have not conducted a formal needs assessment or other research before they began their work. Prevention interventions are often designed and implemented based on staff experience and/or on priorities set in advance by the funding body.

Regardless of the focus of the evaluation project which have been negotiated between researcher and community-based organization, it is imperative to include formative questions as part of the evaluation process. If the focus is set immediately on process or outcome of a specific intervention, the theory or model about the causes of HIV transmission which underlies the prevention work remains unclear. As O'Leary

O'Leary et al. (1997) remark, there is always a theory of some sort operative, even if it has never been articulated. The questions asked as part of a formative evaluation not only clarify the problems which the prevention activities are trying to address but also the organization's understanding of these problems.

Important questions to be asked as part of a formative evaluation include:

- Who has HIV (socio-demographic characteristics)?
- How has the spread of HIV developed over time since the beginning of the epidemic?
- Who is at risk for infection and in what context?
- What is the dynamic driving the epidemic?
- What resources are available for prevention work and has the amount of resources changed over time?
- What organizations are already conducting prevention? What are they doing and for whom?
- What is the unmet need in terms of prevention?
- What strategies would be most effective in changing the course of the epidemic, given available resources?

Data from various sources can be used to answer the above questions. Official surveillance data should be examined as a starting point. Other sources of information can include: the local public health authority; local HIV/AIDS medical specialists; school officials; and representatives of at-risk groups (e.g., gay men, drug users, youth, etc.). Small scale, local KAB surveys and/or interviews with members of at-risk populations (individually or in focus groups) can be initiated to supplement information from the above sources or to provide new information on groups about which little is known. A researcher and the organization for whom he or she is working should not assume that the at-risk groups at the local level are the same as at the national or regional levels, given vast differences in epidemiology based on geography. In gathering data, the researcher should work with the community-based organization to generate hypotheses about who is becoming infected and why so that questions can be formulated more precisely in order to examine these hypotheses. Ideally, staff from the community-based organization will be involved not only in the development of research instruments, but also in the data-gathering process.

The next step is to use the data gathered to construct a model of HIV transmission for the locality in which the community-based organization operates. This model can be informed by theories and practice experience from the literature, from organizations in other localities, from national studies, and from the input of other experts. However, the model developed needs to depict the local dynamic of HIV transmission, which will invariably differ from the characteristics of the national epidemic. For example, to say that young male sex workers are at risk nationally may have no meaning at the local level if there is no sizeable prostitution scene. The model should be complex: dynamic and multi-leveled in structure, but as parsimonious as possible. As with all good models, the idea is not to describe every aspect of the phenomenon, but to include the key elements. From a complexity perspective, this would mean focusing on the elements which appear to drive the overall dynamic of HIV transmission. By initiating this process of model-building based on available data, the researcher helps the organization to expose assumptions which have no empirical basis while incorporating existing theories about HIV transmission into a comprehensive picture.

In a third step, existing prevention efforts are mapped to determine who is doing what and for which population. The services of the community-based organization under evaluation should also be included here, if prevention work is already being conducted. The extent and goals of current prevention programs should be considered. The “mapping” would consist of creating a picture of available prevention services and the populations they reach. (cf. mapping process for prevention services in London as reported by Hartley, et al. 1999).

In a fourth step, the model of HIV transmission and the map of existing services are compared in order to determine current service needs. From a complexity perspective, this is not primarily an exercise in defining the needs of specific population groups. Rather, the focus is on the overall dynamic and identifying at which points in the dynamic interventions are exerting influence. For example, there may be no existing services for gay men in a particular locality. However, surveillance data and other reports show the vast majority of cases among heterosexual women, a trend which has grown over time. The question is not: Which groups are being served and which are not being served? But rather: Which groups need to be served and in what way in order to have the greatest impact on the epidemiological dynamic? In this ex-

ample, strengthening services for gay men may be deemed of little importance, unless it is seen as being integrally connected to the epidemic among heterosexual women. The prevention need is thus to be defined as the need of the system as a whole, which may translate into identifying the needs of specific sub-populations who appear to be of central importance within the transmission patterns.

In a final step, a prevention strategy is developed which is responsive to the identified needs and which takes into account the overall epidemiological dynamic. With reference to complexity theory, this strategy should be characterized by the following:

- A multi-level and multi-factorial approach
- A primary goal of changing the overall dynamic
- Identifying *critical* and *indicator* variables
- An avoidance of “methodism” and an emphasis on appropriate, flexible responses to changing needs

The **multi-level, multi-factorial approach** avoids identifying one factor as primary, and instead includes interventions at various points within the transmission dynamic. The needs of specific populations and individuals should not detract from the **goal of influencing the overall dynamic** so as to prevent the spread of disease in the longer term. **Critical variables** are those which are connected with many others in the system and therefore offer potential leverage for influencing a host of other variables and, ultimately, the system as a whole (e.g., psychosocial characteristics of the target group; unsafe sex behavior). **Indicator variables** are those which are dependent on many other variables but which, themselves, do not exert major influence on the system (e.g., incidence and prevalence). Critical variables can be used for tracking the state of the system over the shorter term, indicator variables over the longer term. Finally, an **avoidance of “methodism”** means remaining focused on the changing epidemic as opposed to being side-tracked into the pursuit of developing exact methods of producing change. A strategy needs to assume that all approaches must be constantly evaluated in order to assess their appropriateness and to be adapted accordingly.

A prevention strategy should include: clearly defined goals; a description of methods to be employed in order to achieve those goals; identified “leverage points” for intervention; as well ways of monitoring progress toward the goals.

**Goals**—Two types of goals should be specified: Those related to critical variables and those related to indicator variables. Goals related to critical variables reflect the immediate effects which specific interventions are trying to achieve, for example (from the model of the prostitution scene presented in Chapter 5): decreasing the psychosocial instability of sex workers, increasing a baseline readiness to engage in safer sex, or reducing discrimination of the prostitution scene. The goals should be defined as concretely and measurable as possible. Goals related to indicator variables reflect changes in the overall transmission dynamic and thus represent the ultimate aim of the entire prevention strategy. Ideally, reliable incidence and prevalence data is available at regular intervals on the target groups. Where this is not the case, other indicators of the underlying dynamic need to be employed (see section 6.4). The achievement of goals should be set within a realistic timeframe.

**Methods**—The methods specified within the prevention strategy are ways in which the critical variables can be influenced in order to achieve a change in the transmission dynamic. The choice of method should be based on practice standards or best practice guidelines, when available. Methods can vary, dependent on the type of variable chosen for the strategy, the type of change being attempted, and the amount of change needed. For example, if the goal is to increase the knowledge of a particular at-risk group (either about HIV or about an issue related to the transmission dynamic), then pedagogical methods of various sorts and/or social learning approaches can be used. If the goal is to address the lack of medical services for those who are HIV positive (and thus decrease infectivity and improve quality of life), then organizational and management approaches are useful and the target group consists of medical providers and public authorities. If lessening discrimination against an at-risk group is a goal, then political approaches are necessary and the target group consists of public authorities, politicians, and the police. Finally, if affecting the baseline probability of unsafe sex among an at-risk group is a goal, then psychological interventions at the individual and/group level would come into question.

**Leverage Points**—The focus in prevention service planning should be on developing an overall strategy. It is this strategy which, in turn, determines which interventions are useful. A collection of interventions without an overall plan is not a strategy; it is simply a collection of interventions. An important aspect of strategic thinking is identifying places within the problem system where change is most possi-

ble (*leverage points*). Once a model for the transmission of HIV has been developed, practical experience and knowledge about the situation of at-risk groups serve to identify what these leverage points are. In the example of the male prostitution scene presented in Chapter 5, the so-called realistic scenario of intervention combinations was conceptualized based on such thinking. Given the great psychosocial need of male sex workers, there is little interest in the practice of safer sex in terms of an abstract goal of remaining healthy. Thus, the baseline probability of unsafe sex is not a leverage point in the system. Addressing the psychosocial need itself is, however, of interest to the target group and also within the means of the projects involved in this work. Thus, creating greater stability in the lives of sex workers is a leverage point in the overall transmission dynamic.

**Monitoring**—An organization can only know to what extent prevention goals are being reached if there is a systematic way of measuring the critical and indicator variables over time. Surveillance data at the regional and local level is indispensable. In addition, other forms of regular feedback from and about target groups are necessary. Information of this sort needs to be gathered in such a way that it complements rather than interferes with daily program operations, which requires identifying methods which can be easily integrated into practice activities. This can include, for example: periodic rapid surveys of target groups; focus groups of community experts; or data collected on project clients. Another aspect of monitoring is to gather information about how the strategy is being implemented, thus providing process data (see 6.3.3).

### 6.3.3 Process Evaluation

A process evaluation is concerned with how a project is conducted. The central question is to what degree the prevention strategy was implemented as planned. Activities are analyzed not in terms of their intended outcome, but rather in terms of the original problem description and the challenges encountered during program operations. The goal of a process evaluation is to clarify which aspects of the strategy are running well and which are not so that improvements in services can be made.

Important questions to be asked during a process evaluation are:

- Is the prevention strategy still appropriate based on current knowledge of the dynamic of HIV transmission?

- Is the strategy feasible given available resources and expertise?
- What organizational or logistical problems have been encountered?
- Are there more efficient ways to carry out the strategy?
- Are the methods used appropriate?
- Is intervention being conducted at the appropriate places (“leverage points”) in the system in order to produce the most change?
- Are all members of the target groups being reached?
- Are monitoring procedures adequate?

The source of process data is the monitoring process set up during the formative evaluation. Data gathered should include such areas as: frequency and extent of interventions; response of target groups; difficulties in carrying out interventions; number of people reached in the target groups; and regular reports regarding critical and indicator variables.

The results of a process evaluation can only improve current services if mechanisms are established for integrating the findings in the decision-making processes of the organization. In this way, the prevention activities undergo an ongoing process of review and adjustment, based on project experience, reactions of the target groups, and the overall dynamic of the epidemic.

#### 6.3.4 Outcome Evaluation

Outcome evaluations are the most commonly reported form of evaluation research and are usually what is implied when the word “evaluation” is used. Here the focus is on the effects of the prevention strategy as implemented. The guiding question is: Did the strategy achieve its goals?

At the heart of the outcome analysis are the indicator and critical variables mentioned above. Through an ongoing process evaluation, a project should be generating monitoring data on a regular basis concerning both types of variables and feeding this information into the decision-making structures. Thus, program services are involved in a process of ongoing improvement and adaptation. An outcome evaluation—often called a summative evaluation—takes place at the end of the timeframe in which the goals of the prevention strategy were to be reached. Thus, all monitoring data for the entire time period is brought together to examine changes which have oc-

curred in the transmission dynamic as a whole as a result of the strategy implementation.

To the degree sufficient data has been provided through the monitoring process, changes in the indicator and critical variables should be observable. During the summative evaluation it may become clear that the data generated is somehow inadequate. In that case, supplemental data may need to be generated post hoc and changes in the monitoring process should be considered which will produce an improved level of information in the future.

However, even with sufficient data and in the presence of observed change, to what degree can this change be attributed to the prevention strategy and to what degree to other influences on HIV transmission in the particular locality? Here the concept of *judicial review* proposed by Keith Tones (1997) is useful, which he defines as “providing evidence which would lead to a jury committing themselves to take action even though 100% proof is not available.” In practice, this means bringing together all the sources of data generated by the monitoring process (with supplemental data, if necessary) and examining the likelihood of the observed changes being the result of the prevention strategy as implemented. The image of a jury implies that such a review is more than the staff of the organization assessing their own work; this should be happening on a regular basis in the context of the process evaluation. A jury is generally a panel of experts who have no direct interest in the matter to be decided and who, on the basis of the information available, judge the particular merits of a case. In this situation, this would mean deciding to what degree the prevention strategy has affected the dynamic of the epidemic. Such a decision involves more than a yes or no answer, but rather an examination of all stages of the strategy—from conception to implementation, a review of the information concerning HIV transmission in the particular locality, and an analysis of the various components of the strategy. Such a jury should thus be able to give detailed feedback to the organization concerning its work within the given timeframe. Just who should sit in such a jury is a central question of both practical and political importance.

Judicial review as the norm for deciding the question of overall program success does not exclude using a variety of methods to analyze the effects of specific interventions. Experimental, quasi-experimental, as well as non-experimental designs can all be deployed, depending on the nature of the questions being asked, the re-

sources available, etc. A range of qualitative methods can also be used. An organization may, for example, include specific methods (e.g. pre- and post-test procedures) as a standard practice in developing new interventions or providing a source of information about intervention effects. Or a judicial review may cast doubt on particular activities which could then be subjected to a more focused analysis.

## **6.4 Case Examples**

To concretize the outlined complexity-based evaluation procedure, it is useful to present two examples. The first is the evaluation of a prevention strategy for male sex workers in Düsseldorf, building on the data presented in Chapter 5 which were used as the basis for the simulation exercise. The second is based on consulting being conducted by the author for a smaller AIDS service organization in North Germany.

### **6.4.1 Prevention for Male Sex Workers in Düsseldorf**

As a result of the needs assessment conducted by the author (Wright in press) a health promotion project for male sex workers is being organized in the city of Düsseldorf. The data gathered within the context of the needs assessment, the other aforementioned studies and practice guidelines, as well as the findings of the simulation presented in Chapter 5 will be used to outline some of the key points of how an evaluation of the prospective project could run. Of course the real process of an evaluation would need to be more detailed and specific than can be included here.

Soon a group of interested organizations will be meeting with public authorities to discuss the findings of the study and to decide what form the project should take. There is already strong support for the recommendation of the author to organize a free-standing project meeting the minimal standards of the AKSD: at least three full-time staff members providing outreach services, a drop-in center, individual counseling, cultural mediation/translation, medical services, networking, and public relations. We will assume here that the group of stakeholders involved in the project will decide to accept this recommendation in full and that the necessary resources will be made available.

The needs assessment, which was based on a triangulation of data from various sources, not only provides specific information about the situation of sex workers in Düsseldorf, but also about how service structures respond to sex workers' needs. Over the course of the planning process for the new project, it will be important to

examine more closely this information in order to create a model of the local epidemic in that population.

One possibility would be the model presented in Chapter 5, which is based not only on data from Düsseldorf, but also on the combined experience of the member projects in AKSD as described in their published guidelines. The underlying concept is structural prevention, as described earlier. The four levels and the accompanying indicator and critical variables were also delineated in Chapter 5. A finding in the needs assessment which was not mentioned above is the very strong presence of non-nationals in the Düsseldorfer scene, which may result in a higher level of risk for the sex worker population as a whole, based on the documented disadvantages experienced by non-nationals. Particularly for this population, but also for sex workers as a whole, the issue of access to existing services would need to be addressed as an important element in the model describing the factors which contribute to a sustained high level of psychosocial need.

A mapping of current health care system structures would reveal a large palette of services for which sex workers are technically eligible, but which go underutilized based on a lack of fit between service structures and sex workers' lifestyles and due to discrimination and a lack of information on the part of helping professionals. A primary role of the new project would thus be to serve as the connecting link between sex workers and the larger structures. For non-nationals the barriers to services are particularly pronounced, a situation which will require special remedial measures.

The initial prevention strategy could consist of the interventions proposed in the "realistic scenario" presented in Chapter 5. This would mean an emphasis on lowering the level of psychosocial need by providing basic social work and medical services. In addition, the infectiousness of HIV positive sex workers would be lowered by encouraging them to come forward and receive medical care as well as to consider ways to lower the probability of transmission even when engaging higher-risk sex. Activities to promote safer sex as well as to reduce discrimination against the prostitution scene would also be conducted. For each of these levels of intervention, appropriate methods need to be identified, based on the experience of the projects and of practitioners in Düsseldorf. Also, goals need to be set for a specific time-frame, for example, for the first three years of project operations, and ways to monitor progress need to be developed.

In the simulation exercise in Chapter 5 we were able simply to set goals by changing percentages and running the model, monitoring progress by reading agent values. For the real life project in Düsseldorf it will be more challenging to identify concrete and measurable markers of change for the critical variables. Taking the variable psychosocial instability as an example, we see that this problem is not insurmountable. To address the needs of the sex workers we have the project itself. We also have a network of services for which the sex workers qualify, but which they do not access. Therefore, two target groups could be identified for interventions conducted by the project: the sex workers themselves and the staff of programs in the social welfare system. For the sex workers, the interventions would be: the drop-in center, counseling, and medical services. For program staff they could consist of in-service trainings on the needs of sex workers. The level of need and the utilization of other services as reported in the needs assessment could be taken as the baseline level. This could be supplemented with ongoing information provided by clients at the time of first contact with the project. The client need could then be surveyed at regular intervals. This would be supplemented by information on the utilization of services elsewhere by asking the clients at regular intervals where they are clients and what they use as well as from statistics provided by the projects. Concrete goals for the target group clients could be, for example, a 25% reduction in homelessness. For the social service providers a goal could be increasing sex workers on their case loads by 30%. Qualitative goals would also be possible, for example, asking sex workers about the level of discrimination at the various social service agencies at the beginning of the three years and asking them again mid-way through the period and again at the end.

Identifying indicator variables to examine the structure of the epidemic as a whole poses a greater challenge. Generally available surveillance data is not of any use, given that sex work is not reported as a category. Including HIV status as part of the general questions asked of clients would be a way to monitor at least a portion of the population. Rapid surveys, as conducted during the needs assessment, provide another option. Working with the local testing site to monitor the use of services by sex workers and the proportion infected would be another source of data. The fact that project estimates (Wright 2000b) and observed levels of HIV prevalence (Wright in press) could be so close provides support for the use of staff estimates as an addi-

tional source of data. These various sources could be triangulated to produce a picture of the epidemic at a point in time. The focus would be on gauging trends in the epidemic as a whole over time, rather than on determining an exact point prevalence.

The process evaluation would consist of interpreting the data generated by the monitoring system at regular intervals. For example, the information could be summarized in reports and distributed to staff to be discussed at team meetings. Identified problems could be discussed and solutions proposed. Supervision could be another forum for resolving conflicts revealed in the process evaluation.

At the end of the three-year period, all monitoring information would be summarized and reviewed by a “jury.” The AKSD is a ready-made group of experts suited to this purpose. The AKSD could be expanded with one or two experts from outside to provide additional perspectives.

#### 6.4.2 Prevention in a Small North German City

Over the last year the author has been providing consultation to an AIDS service organization in a small North German city. The focus has been on re-defining the services based on the needs of residents.

Initially, the author and a staff member from the Deutsche AIDS-Hilfe were asked to conduct a weekend workshop for all members of the organization (paid staff, volunteers, and board members). The focus of the meeting was to be the discussion of a new concept paper for the work of the organization as a whole. A turnover in staff and board provided the opportunity to re-examine the organization’s mission and goals. Over the course of the weekend it became clear that the services mandated by the public health authority had never been examined in terms of the actual needs of the city’s residents. For over ten years the organization had been funded primarily to provide services to people with HIV/AIDS, and this in a part of the country where the prevalence of HIV has remained low. The epidemiological situation combined with the experience of the organization suggested that primary prevention needed to be a focus of their work. They had, however, never been funded to do such work. Thus, the agreement negotiated between the author and the organization was to work on developing a primary prevention strategy for the city and to define the organization’s activities accordingly. On the basis of the initial consultation, the organization’s

board managed to re-negotiate a portion of their funding to include prevention activities. This freed up staff time to concentrate on developing an appropriate strategy.

A follow-up consultation focused on helping the organization to design a needs assessment for the city. This includes expert interviews (from schools, public authorities, social service and medical providers, and representatives of potentially at-risk groups); focus groups with potentially at-risk populations; a review of regional surveillance data; supplementary data on the course of the epidemic from the local public health authority and HIV medical specialists; and KAB surveys of gay men and youth. Both information about current prevention services as well as about potential levels of risk are being gathered. These results will be reviewed in order to map current services and to produce a model of the epidemic.

Two challenges are faced by this organization which may be typical for many AIDS service organizations in Germany. The first is the aforementioned contract with the public health authority which specified tasks having no basis in documented need. The second is a shift from priorities set at the national level by the Deutsche AIDS-Hilfe which, for both political and epidemiological reasons, has concentrated on the needs of particular at-risk groups. To what degree these priorities reflect the epidemic in this small city has never been examined.

The structure of the epidemic has likely remained stable, with emergent problems in terms of risk (lack of knowledge, etc.) which may be identified among certain groups. Such changes would not be reflected in surveillance data, given the low baseline prevalence. Thus, the model of the epidemic which is likely to emerge will be one of little or no change. The general goal will likely be to keep prevalence at its current level and the structure of the epidemic stable so that no new groups become infected. In several respects, this presents a simpler picture than infection among male sex workers where there are several potential leverage points in the dynamic landscape of contact. In complexity terms, the epidemic has remained at a low level of complexity, given the small number of people involved. The model should thus be more straight-forward, depicting less variables and target groups for consideration. The difficulty in such a situation, however, is identifying the likely route that the epidemic could take should the pattern change and the number of new cases begin to rise. The challenge is to identify potential new pathways which the epidemic may follow and to intervene accordingly.

### **6.5 Evaluation as a Complex Process**

The application of complexity theory to evaluation research makes clear the complex structure of the evaluation process itself.

The formative, process, and outcome aspects of an evaluation are not three independent activities conducted in a linear fashion, one after the other, but a nested hierarchy of levels comprising the overall evaluation process. As Figure 20 suggests, the formative evaluation is the framework within which the other two levels function, thus representing the prerequisites for the process and outcome evaluations. The process evaluation has a mediating functioning between the concepts and goals outlined during the formative evaluation and the end effects of the strategy. Finally, the outcome evaluation is concerned with the extent to which the goals were reached. Typically, evaluation research concentrates on outcome evaluation—particularly certain aspects of the outcome evaluation, such as the effects of specific interventions—and thus considers only the proximal causes of program success or failure. As we saw most clearly within the course of the social simulation in Chapter 5, a focus on proximal causes fails to exploit the potential of harnessing the transformative power of interventions at other levels. In terms of evaluation, this means providing research support for both the formative and process aspects of the work so that the interventions which are finally executed and measured have a clear empirical connection to the problem which they are attempting to address.

Other important aspects of the complex nature of evaluation research include feedback loops and the open quality of the evaluation system. The feedback loops are built into each of the three levels in order to provide a constant stream of information input into a project regarding the problem at hand, the target groups to be reached, and the operating of the program itself. Thus, evaluation is not a static yes-or-no oriented pursuit to judge the appropriateness of certain actions, but rather an ongoing process which supports the continued development within an organization. The dynamic quality of the evaluation process is also evident in the open nature of evaluation research activities themselves. Over time, the techniques used to evaluate the work of the organization will themselves evolve, as the understanding of the problems deepen, experience with monitoring procedures grows, and initial models and hypotheses coalesce into theories describing the dynamic of the local epidemic (cf. process of quality management in ISO 2001).

From a complexity perspective, the current debate about evaluation methods and evidence is the result of neither the scope of HIV prevention nor the accompanying evaluation research being adequately taken into account. When the complexity of both is recognized in the formal sense, then the focus can shift to developing approaches which take on the larger issue of how to produce useful information over the longer term so that the community-based organizations as a whole can continue developing in expertise and quality.

## **6.6 Evaluation and Quality Assurance**

The further development of health care structures in industrialized countries has focused on identifying and strengthening those elements which contribute to *quality of care*. The discussion of quality in this sense has evolved from management theory and practice which has produced an enormous literature on *quality assurance*, *quality control*, and *quality management*. The overriding concern in this literature is to describe the characteristics of organizations which will result in a consistently high level of product quality, which is seen as being related to such factors as cost-efficiency, consumer satisfaction, etc. Quality management for healthcare and social service agencies has emerged as a field unto itself, applying the theory and practice of private sector management to these sectors (cf. Altgeld et al. 2000; Ruckstuhl et al. 1998).

The evaluation literature on health care interventions makes little reference to the discussion of quality assurance, because evaluation practice existed prior to this discussion and because the focus of evaluation is somewhat different. Evaluation research as represented in the scientific literature is predominantly concerned with relating specific interventions to particular health outcomes. As such, evaluation research creates a link between what organizations in the health care system do and the larger epidemiological realities of the world “out there.” This link is of crucial importance for long term health policy planning.

Conceivably, an organization demonstrating high standards of quality may be engaging in practices which do little to address the given health problem at the population level. As we saw in the example of the AIDS service organization in the small North German city, services were being provided for years to people with HIV. As long as those living with the disease are seen as being the primary consumer group,

and given that certain quality management standards are being maintained, such an organization could be providing exemplary care. However, from a public health perspective, even such a high quality of services would be a problem. As discussed above, in a region of low prevalence, the public health priority needs to be placed on primary prevention and thus disease containment at the population level. This connection to the larger epidemic is only possible by supplementing a quality management perspective with an ongoing research-based evaluation process.

What is difficult to imagine is the opposite of the above: Namely, an organization successfully implementing an ongoing evaluation without having an understanding of quality assurance. As the description of the various aspects of the evaluation process make clear, the prerequisite for an organization being able to conduct an evaluation is a cohesive and well-operating structure and a level of organizational motivation sufficient to support such a self-reflective process. Of course, the standards of private industry cannot be applied to the non-profit sector. However, management principles have been developed to enable social service and health care organizations to function with higher levels of efficiency and with improved staff and client satisfaction. This includes processes for goal setting, accountability, decision-making, and project planning and implementation (cf. ISO 9004 2004). Evaluation research can therefore be seen as part of an overall quality assurance strategy within an organization working in public health. The information generated not only provides input for internal processes of organization improvement, but also constructs a bridge between the larger health care issues being addressed and the mission and structures of the organization (cf. role of data-analysis in decision making as discussed in ISO 9004 2001).

## **6.7 Conclusion**

In this chapter we drew together the observations from all previous chapters in the interest of presenting an approach to the evaluation of community-based prevention which takes into account the complex dynamics of an HIV epidemic. Building on the conventional division of evaluation into formative, process, and outcome tasks, a complexity-based evaluation can be defined as the practice of describing and responding to the dynamics of the epidemic at the local level. The primary focus of prevention is accordingly placed on articulating a strategy to influence the larger

processes fueling the spread of HIV. Evaluation is thus concerned with guiding the conception of such a strategy, providing feedback as to how the strategy is carried out, and monitoring indicators of the strategy's influence. From a complexity perspective, the effectiveness of individual interventions at a specific level of the epidemic remains of concern, but only within the context of the overall strategy. In this way, a complexity-based evaluation keeps us focused on the bigger picture, prompting us to gather information about a strategy's performance from various sources and to interpret this information judiciously based on what is known about the various variables driving the epidemic and their interrelationship.