The Role of Sociology in the Study of Mental Health
... and the Role of Mental Health in the Study of Sociology*

BLAIR WHEATON
University of Toronto


This essay considers the dual roles occupied by the sociologist of mental health. These roles involve the articulation of social causation in the study of mental health outside of the discipline, coupled with the articulation within the discipline of the importance of mental health in the study of sociology. I consider these roles both through examples and speculation, emphasizing the unique combination of conceptual and methodological tools that define the intellectual terrain of this area of sociology. The advantage of this dual role—of looking outward while also looking inward—is that we are able to draw from the essential developments and innovations from one source and “move” these insights toward the other. The difficulties of this position are also clear: As an area, we may be structurally marginal from both perspectives, at the same time that we offer considerable analytic power that could significantly impact the direction of research involving mental health in both realms.

I remember very early in my career the comments resulting from the first paper I submitted to American Sociological Review. Along with the mainly friendly suggestions, there was advice: It was too early in my career to offer opinions about the needed directions the literature on mental health should take in the future. So I have waited, and according to that edict, it’s now or never.

Honors are a rare event in the academic life course, but I must begin by emphasizing that receiving an award named for Len Pearlin, especially as the first recipient, is an honor with specific personal meaning to me. This occasion is also an opportunity for me to express what must be interpreted as feelings, more than judgments, about the collective role of the sociology of mental health, both within our own discipline and in the allied disciplines that also target mental health as an issue. As sociologists, we persist in asserting that sociology has a crucial role to play in understanding mental health processes. The question is, are enough people listening? I have a great deal of respect for sociologists who work in this area, making an argument, on the one hand, to the medical and biological professions about the role of social environment, and, on the other, making a case to the rest of sociology about the integral role of emotions, of the inequality of the distribution of emotion, and the effects of that inequality on social life. This dual role is both a difficulty and, in some ways, our greatest resource as a substantive area. It is both the contributions we can make as an area and the challenges we face that I want to focus on in this essay.

I start from the assumption that we have two interrelated basic claims. First, social environment is both a starting point and a carrier of the fact of emotional inequality, and second, variation in emotional functioning has specific social consequences for our institutions and

* This essay was first presented as a talk prepared for the acceptance of the Leonard I. Pearlin Award for Distinguished Contributions to the Sociology of Mental Health, given by the Mental Health section of the American Sociological Association, August 15, 2000. This revision for publication benefited from the helpful comments of three reviewers for JHSB. I have retained some of the informal quality of a talk in this final version. Direct correspondence to: Blair Wheaton, Institute for Human Development, Life Course, and Aging, 222 College St., Suite 106, University of Toronto, Toronto, Ontario M5T 3J1.
our society. While these claims may sound familiar, the prevailing public consciousness at this point in history, as well as within some sectors of the medical and social sciences, often seems to forget, if not ignore, these basic facts. For example, I am forever struck by the historical reach of what I call biomania—the implicit belief that if we just drill down far enough in biology and genetics, we will explain most of the variance in most behaviors. This is not just an issue within granting agencies, it is in the public mind. Listen to your aunt or your sister or your father discuss why you are as you are, and somewhere, inevitably, much of what you do will be mapped post-hoc to some familial genetic imperative. A conversation I had at dinner just last night at this meeting is instructive. As if to anticipate the subject of this talk, a colleague in another area of sociology scoffed at my interpretation of a common friend's heart attack as in part due to a particularly stressful period of his life, claiming that public pressure about his work wasn't "real stress," and that the heart attack was more likely to be due to "constitutional factors." This interpretation may indeed be plausible, but I was struck by how quickly biology was the preferred alternative hypothesis. In fact, when we look closely at what research suggests, what we find is that the uncertainty in biological or genetic explanation is usually explicitly admitted in the conclusions of researchers themselves, e.g., in the form of statements that the heritability or explained variance is still only 50 percent. However, this fact somehow gets lost in the social, political, and communication processes that turn complexity into simplicity, continuum into category, and thus probabilistic reality into assumed certainty.

I remember a National Institute of Mental Health program person who came to our review committee in the late 1980's and reported on the distribution of funding that year for the study of schizophrenia. Twenty-four out of twenty-five funded grants were about biological determination. One was about social processes in treatment. I asked whether anyone at NIMH was aware of Link, Dohrenwend, and Skodol's (1986) then recent article on a kind of re-invention of the social causation of schizophrenia, and the response, literally, was, "well, no, we don't do that much anymore around here." The often-cited view that the best strategy for investment in science involves diversification seems not to be reflected by this response. Of course, we know that "what goes around, comes around" is also the case, and so we expect and see signs of the re-emergence of attention to the social environment, and more generally, to social experience.

However, such changes are relative, and a question still needs to be asked: Does biology deserve that much emphasis in the public discourse concerning mental disorder? If you believe, as a social scientist, that social environments often act as the starting point for the requisite biological changes implicated ultimately in mental health changes, and you also understand that indirect causation is still causation (Hume 1739; Bunge 1959), then your answer is likely to be "probably not." The most common role articulated for social causation in conjunction with biological causation imagines a primarily synergistic relationship, in which social environment acts as a modifier or amplifier of biological processes. This kind of interactive role looks beyond a more basic possibility: that social environment produces the biological change that in turn acts simply as a more proximal cause of change in mental health. This role needs to be addressed more broadly in future research, both because there is evidence in support of it (Fremont and Bird 1999), and because it points to social environment as a "fundamental cause" of mental health problems (Link and Phelan 1995). At the same time, we should remember that biological mediation is not a requirement for the establishment of social causation, just as structural equation models make clear that removal of a mediator actually illuminates the total causal effect of a variable.

SOCIOLOGY IN THE STUDY OF MENTAL HEALTH

If I think about what sets the sociology of mental health apart from allied disciplines that study mental health, I gravitate to two issues. First, as a group of researchers we operate ultimately as a basic, and not an applied, science. This is a good thing, I believe. The immediacy of the demand for intervention can at times result in acceptance of the current state of the art as a "given." This produces an inertial component in research that not only may be misleading but may also delay better explanations. Does a life event list measure all that is stressful? We know this is not true. The problems of
measurement should not prevent us from incorporating important concepts into our thinking while we also address those measurement issues. Remembering that the concept (the "stress universe") is separate from the measure (life events) is more easily achieved without the immediate pressure to intervene or help—at least in the short-run. Moreover, in the long run, intervention will surely benefit from a broader conceptual and measurement approach to stress, including both acute and chronic, macro and micro, and childhood and current sources of stress. Ultimately, basic science can play the important role of correcting the problems attending the automatic application of popular and prevalent ideas that may in turn impair the effectiveness of applied work by re-defining the problem, the model, the question, or the explanatory framework.

It is easy to find examples in the work of sociologists of mental health pointing to fundamental changes in direction or redefinitions of the issue at hand, made possible, I believe, by the benefits of distance from prevailing assumptions and the need to intervene. The stress process model is an obvious, and in the context of this talk, directly pertinent example (Pearlin et al. 1981), but there are numerous other ideas emanating from the sociology of mental health that have grown into standard operating models or assumptions and yet countered the prevalent practices at the time. Thus, we now think naturally in terms of the multiple outcomes of stress (Aneshensel et al. 1991), of the social embeddedness of stress (Pearlin 1989; Turner, Wheaton, and Lloyd 1995), and of the labeling process as not inherently at odds with the "psychiatric" model but rather the two approaches having dynamic and complementary roles to play in a larger explanation (Link 1987). Some ideas, for example, emphasizing the inherently dimensional nature of mental health problems (Wheaton 1982; Mirowsky and Ross 1989; Kessler 2000) still await widespread acceptance—or understanding. In metaphoric terms, our role is to be able to say, "sorry, we are on the wrong road, we need to back up to that last fork and take the other road."

Second, sociologists occupy a unique position in terms of the range of methodological techniques at their disposal. While sociology is rarely a source of innovation for new techniques of analysis, I suspect it is the most diverse training ground for methods in the social, behavioral, or health sciences. True to its structurally marginal position in the presence of more funded disciplines, and its recent history, it has been oddly blessed by the absence of a single methodological paradigm or the focused content of its subject matter. Psychology as a discipline lives and breathes the experimental model as a reference point. Economics has a currency that acts as a visible basis for all human exchange. The health professions have a single issue on the table, albeit a big one, but most certainly the way in which research is planned and designed is a response to the very nature of the one dependent variable at issue: health vs. illness. In the context of more focused agendas, narrow band methodologies feed assumptions that loom large. Assumptions insidiously become unassailable truths: They are repeated as premises over and over until they become so self-evident there is no need for explicit reference. In sociology, we seem to absorb practically everyone else's innovations and then nurture them into elaborated techniques which come closer to meeting the inherent complexity of the data and design situations that we regularly face in this discipline. Again, this is a good thing.

There are cultural aspects to the dominant research assumptions in both psychology and epidemiology which are questionable but survive. I once counted 15 introductory psychology texts on the same university library shelf that somewhere in Chapter 1 dismissed non-experimental research in general by referring to it globally as "correlational." In sociology, and ironically in psychology itself, the evolution and generalization of methods for analyzing panel data suggest the assumption that causality can only be investigated in experiments is both misleading and limiting. The impact of thinking in terms of structural equation models, rather than just individual equations, is still uncertain in disciplines such as epidemiology (and indeed in parts of sociology), primarily because the general role of the assumptions in these models is widely misinterpreted. An attractive feature of structural equation methods is precisely the explicit and visible nature of the assumptions used. When we forego the clear representation of both the measurement and structural specification assumptions embodied in latent variable structural equation models for the apparent simplicity, straightforwardness, and safety of standard regression equations, what we really achieve is
greater ambiguity and uncertainty about how to interpret results.

To be specific about this, the most common result of the treatment of predictors in regression equations as “co-equals” is confusion about the role of indirect effects. This is the case because some variables in the equation are almost certainly going to be mediators for the effects of others, but the effects we see in the coefficients are only “direct effects.” This implies that the total causal impacts of variables, including both direct and indirect effects, are partially hidden by the equation. Furthermore, the interpretation of direct effects mistakenly as total effects implies that all other variables in the equation are background controls rather than mediators. To cite an actual example of the problem, consider a paper on the effects of single parent status on health taken from an epidemiological journal. I do not intend to suggest in using this example that the problem exists primarily or only in epidemiology, and to make that point I prefer not to cite the source specifically, which could imply unfairly that this paper is somehow unusual. This paper interprets household income as a confounder of the effect of single parent status. The fact that household income is also changed by single parent status and thus is in part a mediator, perhaps predominantly a mediator, is not considered. The direct consequence of this is an underestimation of the impact of single parent status, and thus potentially, the unintended misdirection of a research literature.

The unique feature of sociology here may be the hybrid evolution of its methodologies, itself a function of its structural marginalism in the realm of health disciplines. I maintain that the “typical” sociologist has been made more regularly aware of the interpretive problems attending dominant analytical approaches in the last quarter century than colleagues in closely allied disciplines. In these disciplines, the existence of stable methodological paradigms over longer periods of time may indirectly be the reason for inattention to innovation beyond the boundaries of these paradigms—a case where strength eventually breeds vulnerability. In other words, sociologists have the ability to use the methodological terrain of their own discipline to clarify and inform issues raised in other disciplines, and this is an essential part of what we have to offer.

The Category vs. Continuum Issue

Whether we realize it or not, sociologists of mental health may also have unique perspectives on the relationship between concepts and the methods used to study them. Sociologists are more naturally comfortable with continuous variables than certainly the health professions, who must make decisions about the need for intervention, and who have derived the concept of mental disorder from the physical illness model. Over twenty years ago I co-taught an interdisciplinary sociology of health seminar at Yale which regularly included discussion of the history of the development of mental disorder as a categorical concept. In these discussions, I usually argued that clinical history could not possibly produce an optimal arrangement of symptoms into disorders, primarily because of the problems resulting from the use of patients as a sample of all cases of a given disorder, and that the evidence necessary to support the concept of disorder as a category (i.e., an illness) could not possibly follow from building that assumption into current approaches to diagnostic measurement (e.g., using screening questions). I also argued that the conceptualization of mental health as a continuum and as a category were not inconsistent, and that one need not have the same concept for studying causation as for deciding who to treat, although a known relationship should exist between the two. Thankfully, at least two of the students in that seminar found the argument interesting enough to pursue further (Mirowsky and Ross 1989).

Ultimately, I have no a priori objection to the specification of mental health problems as a category, but I would feel more comfortable if I knew that the decision about the existence of the category followed from findings about the continuum (Kessler 2000), rather than the dominance of a theory about a category disqualifying the relevance of a continuum. My concern was, and still is, that pre-categorization may impair optimal understanding of mental health problems, and in some cases shift our understanding of the role of risk factors.

When we consider the use of screening questions in diagnostic interviews as an example, we see how subtle the problem may be. Screening questions may make otherwise long interviews practical, while defining necessary symptoms for the occurrence of disorder, but
they also skew the ensuing distribution of symptoms in that category. If I have never had either of the two specific screening symptoms for depression, I will not be asked the long list of other depression symptoms. Even though I may have had many of those symptoms, my score on any depression measure, using the assumptions of the diagnostic approach, will necessarily be zero. The problem is that the definition of the disorder requires gateway symptoms, and this assumption is typically built into diagnostic approaches rather than evaluated as a part of the research process. As a result, we cannot discover the actual distribution of depression in the population, and we cannot discover how much depression content may overlap with the content of other discretely defined disorders.

In considering how to frame our concepts, we should also consider how our methods work conceptually. Even methods for categorical outcomes like disorder, for example, as exemplified by log-linear or logistic models, have to transform that outcome into a continuum to interpret results. In the end, we always end up interpreting in terms of some continuous concept, such as odds, probabilities, or hazards. This is not really just a technical issue. It is a metaphysical clue about what it takes to even conceive of multiple causation and to allocate roles to potential explanatory factors. The fundamental reality is that in order to assign specific roles to multiple factors in an explanation, in which some may be central and others may be peripheral, we need a continuum, because only a continuum allows specification of partial determination and relative weighting of factors.

The derived argument that accompanies the categorical imperative is, of course, that the common everyday scales used to measure depression or distress are tapping into "non-clinical" and probably self-limiting problems. This argument has had unfortunate consequences for the sociology of mental health historically, which has more typically relied on distress scales (while also incorporating diagnostic measures in recent years). Findings about distress can be easily dismissed as having little to do with serious, stable, and consequential mental health problems.

My view of this argument has always been that it is dead wrong. The argument proceeds from a misunderstanding of the difference in nature of categorical vs. continuous concepts. There are two related questions: (1) whether the upper regions of distress scales enter the content realm of serious self-perpetuating problems and (2) whether equal differences on distress across the whole scale stand for equally important variation in mental health. These issues can be assessed but have tended to be couched in terms of using distress for screening purposes, not as a proxy for mental disorder.

Again going back over twenty years, I was interested in the problem of fit between what is measured by distress vs. diagnostic classifications. Table 1 shows the results of some logistic fitting of the Langner Symptom Scale (Langner, 1962) as a predictor of a depression diagnosis using the SADS (Schedule for Affective Disorders and Schizophrenia) in the New Haven community survey from 1975–1979. I use this example because it uses a diagnostic instrument (the SADS) that allows more clearly than later more standardized approaches for the kinds of latitude and speci-

### Table 1. A Classification Table for Predicting SADS Depression From Langner Symptom Scores (New Haven Data).

<table>
<thead>
<tr>
<th>Actual SADS Diagnosis:</th>
<th>Predicted Absent</th>
<th>Predicted Present</th>
<th>Total</th>
</tr>
</thead>
<tbody>
<tr>
<td>Absent</td>
<td>395</td>
<td>41</td>
<td>436</td>
</tr>
<tr>
<td>Present</td>
<td>4</td>
<td>31</td>
<td>35</td>
</tr>
<tr>
<td><strong>Total</strong></td>
<td><strong>399</strong></td>
<td><strong>72</strong></td>
<td><strong>471</strong></td>
</tr>
</tbody>
</table>

- Sensitivity = 31/35 = 89%
- Specificity = 395/436 = 91%
- False Positives = 41/436 = 9%
- False Negatives = 4/35 = 11%
- Positive Predictive Value = 31/72 = 43%
- Negative Predictive Value = 395/399 = 99%
ficity often seen as essential in making clinical diagnoses, and thus may act as a more believable “gold standard” for some. These results are not based on a simple chosen threshold for the Langner scale; rather, they are based on a logistic regression of the SADS diagnosis on the individual Langner symptoms, thus allowing for optimal weighting of symptoms in predicting the diagnosis. The issue of replication of such weights is always an issue, but a side issue here: The logic of protesting on grounds of replicability depends on the accumulation of similar published results, which remains an open question. The symptom-weighted predictor of the probability of a diagnosis is divided at a predicted probability of .075, based on the actual prevalence of SADS major and minor depression in the sample.

How one views this table depends on one’s training, research goals, and assumed professional mandates. A very simplistic rendering, but not at all unfair, would be to say that what you see in this table depends on whether you see the glass as 90 percent full, or 10 percent empty. It is quite clear that the Langner symptom composite approximates 90 percent of the case dispositions in the sample correctly. From the point of view of using the Langner scale as a screening tool, or as a basis for need for treatment, one can validly point out that it still misses 10 percent of the cases. From the point of view of studying causation, however, it seems overwhelmingly obvious that these two measures are not measuring qualitatively different things. In other words, if we were to use the Langner scale composite as a mental health outcome, are we to claim that the results have no implications for “real” depression? The issue is pivotal, because even within the social sciences, there are continuing attempts to keep distress and disorder in different conceptual universes. I believe the necessary evidence for this assumption does not exist.

There are other ways of rendering the relationship depicted in Table 1. For example, qualifying on the Langner scale multiplies the odds of being a “true” case of depression almost 75 times—an odds ratio quite beyond the realm of what we observe for purely substantive (as opposed to measurement) relationships. The phi coefficient (the correlation) here is .58, which could be interpreted as a decent factor loading if the diagnosis was the “true score” and the Langner classification was the “observed score.” However, this correlation does not tell the whole story, since the upper value of phi is affected by the marginal distributions of the variables, and here the maximum possible value is .675. Thus, the adjusted phi is actually .865 (Bohrnstedt and Knoke 1994), suggesting a high level of correspondence as measures. Although this is a very hypothetical number, it is also appropriate to say that the unadjusted phi could be even more misleading.

There is nothing unique about the Langner Index as the basis of these findings. In fact, I also used best approximations from the data to sets of CES-D symptoms, SCL-90 depression and anxiety sub-scales, and the Gurin Index, as well as a best overall predictor from all scales. Interestingly, all results were within a few cases of the result in Table 1. It is worth noting, therefore, that dedicated depression scales did not do any better than mixed depression and anxiety scales, like the Langner Index, in predicting depression diagnosis.

There is another issue in considering distress as a concept that is directly informed by methodological tendencies across disciplines. The assumption that distress only measures moderate mental health problems rests in part on its presumed instability over time. Dohrenwend and Dohrenwend (1969) reported, for example, that many of the individual symptoms typical of distress scales showed considerable change over time. This evidence is two layers of generality away from the appropriate evidence for or against stability. First, we would expect more change on individual symptoms than on an index of symptoms, in part because of greater random measurement error in the individual items, and in part because we should expect some functional interchangeability of symptoms, allowing for the abatement of one problem while another related problem appears. It is the totality of the phenomenon that measures the stability of distress, not patterns of change in specific symptoms.

Second, an index of symptoms still contains measurement error, and thus underestimates true stability. True stability is best estimated by the correlation of latent variables for distress symptoms measured over time in panel data (Wheaton et al. 1977). In case one prefers to be skeptical that such “true” latent variables exist, one should remember that the belief that there is error in measurement, a concept with widespread acceptance, logically implies that a
“true score” must exist: The error has to be error from something! Or, to put the same observation in Bollen’s (1989) terms: The true score is “what remains when the errors of measurement are removed” (p. 208). In preparing for this talk, I reviewed estimated stabilities of distress in LISREL models I have estimated over the last 25 years across 8 different panel data sets, most of which are well-known and span 30 years of time. The basic conclusion is that these stabilities varied from a low of .42 to a high of .77, depending on the length of the lag measured (between 2 and 10 years). One can consider this a moderate level of stability through time. My view of this and earlier results is that distress is not transient and self-limiting; findings such as those in Table 1 reveal a high amount of overlap with more clinical measures; and thus distress taps into important and stable differences in suffering in the population that are, nonetheless, amenable to intervention. This result also suggests that the concept of distress should be useful for studying both structural sources of emotional inequality and the ways in which mental health responds to changes in the life course.

MENTAL HEALTH IN THE STUDY OF SOCIOLOGY

On the other side of our dual role, we are responsible for making mental health an essential concept across the other areas of our own discipline—from gender to family to work to stratification to social change to networks to organizations to social psychology to demography to theory to fill in your favorite related interest in sociology. I am still struck, after over 20 years of talking to colleagues in other areas about social variation in mental health, how often they see little relevance for the explicit consideration of mental health in the work they do. I interpret this as a situation that is partially our collective fault, and perhaps a reflection of a collective blind spot about the importance of studying social selection as much as social causation.

One of the problems we face is the concept of mental health itself—or more precisely, mental health problems. Obviously, a great deal of measurement work has been conducted in the last quarter century in other disciplines to establish a well-developed realm of content for mental health, certainly as much if not more measurement work than in most of the traditional core areas of sociology. However, progress is slow within the discipline in recognizing the richness and variety of content in the concept of mental health. Rather, the reference point in the realm of subjective outcomes involves a small set of simple questions used to indicate happiness or well-being. What is often not realized is that these measures, by virtue of their simplicity and their content focus, are not sufficiently tapping the content on the negative side of the continuum, which may be the defining realm for describing stable differences in the emotional content of lives. In saying this, I am not requiring that positive and negative affect operate as separate factors; rather, the difference may be in the stability of positive vs. negative affect and therefore the weight of their importance in pinpointing essential differences in the overall subjective quality of life.

The implication of this issue is this: We may be able to get to the heart of what varies across social position and social location more clearly by articulating the variation in negative affect than the volatile and elusive issues in positive affect. There is little doubt in my mind that we also measure the negative variation better: In getting down to the molecules of the basic matter of mental health by asking many narrow band questions on our better scales and diagnostic interviews, we circumvent the demand characteristics implied by the “big questions” used to tap positive states of happiness and satisfaction. A prevailing, but I believe, ironic assumption is that we ought to strive to maximize good, and if that means maximizing well-being, then we may be taking on a relatively impossible task. Rather, the world may benefit more if we simply target the reduction of harm. This in itself would be a significant accomplishment, and the measure of this accomplishment would be how much we have reduced the various indicators of suffering we call mental health problems.

Articulating Social Causation and Social Selection

Although we understand that our responsibility even within sociology is to demonstrate the role of social causation in mental health, as part of a larger tradition concerned with specifying the linkage between social structure and the individual, I sometimes think we do not
underscore clearly the direct relevance of this fact to the work of our colleagues. I often accuse colleagues in sociology of a sort of false consciousness about mental health. Many of them, frankly, feel safe from such seemingly ephemeral concerns. What I try to claim is this: realized or not, mental health is everyone’s ultimate dependent variable. This is only a bit of stretch, in fact. I try to point out that an essential source, or reason, for their concern about the increasing prevalence of divorce, or crime, or unequal access to opportunity, or distribution of income by gender, or differential rates of social mobility, or increased disruption of employment in a work career, or work-family conflict, or globalization and standardization of culture and business, or the need for regulation of harassment in the workplace, or probably the next three examples you can think of, is the ultimate effect of these issues on the mental health of those in the disadvantaged position—the state of well-being, suffering, misery, or hope people in these social situations face as a result of the explicit issues raised. In other words, it is important to insist that the implicit assumptions about mental health driving the research questions in other areas be tested explicitly. This will often involve testing hypotheses about the social causation of mental health—an issue that should regularly surface in the sociological landscape.

On the social selection side, the point is simple: As an area we need to make this case more comprehensively and more clearly within sociology. After all, the implication is that our most fundamental models for core issues such as the transmission of social inequality may be wrong, to the degree that they exclude mental health in the role of an independent variable in the determination of social outcomes. In the last twenty years, we have done much better than formerly in demonstrating the net impact of variation in mental health on a range of educational, job, and marital outcomes (Link 1982; Kessler et al. 1995), but the broader implications of these findings for literatures in other areas require further attention. The essential question that still needs to be assessed is whether the impact of emotional inequality on social inequality is independent of, or spurious due to, the impact of more basic social causation. In effect, if we want to argue that mental health matters to social and life outcomes, we cannot simply assume that influence independent of social background has been already demonstrated.

The same argument can be extended to apply to the theories we use to explain mental health variation. Stress theory is not just about differences in mental health; it applies potentially to a wide array of life outcomes (Aneshensel 1996). The important point to make in the context of influencing the discourse of our own discipline is that stress exposure can influence social destinations, whether or not this exposure manifests itself as mental health problems (Wheaton 2001). This argument follows especially from consideration of the effects of childhood traumatic stress on later life, where the process of dealing with stress during formative or transitional life stages may compromise the ability to attend to other contemporaneous life tasks.

CHALLENGES AND POSSIBILITIES

We also face a number of challenges that will determine the role we play in the explanation of mental health processes in the future. I would like to focus on just two of these challenges: (1) learning to evaluate evidence more dispassionately, and (2) pursuing ideas when supportive evidence may be particularly complex or elusive.

Evaluating Evidence

I sense that there are important barriers to the appropriate interpretation and use of evidence, and the consequence is, unfortunately, that we are not able to progress beyond faulty assumptions about the empirical world as efficiently as we must. Whole generations of researchers can make assumptions about “what is” that are indeed questionable. Why does this happen? The most important barrier, I believe, is our own beliefs, although biases, power, politics, and temporal precedence all create noise in the process of empirical clarification. There is a fundamental tension between experience and evidence that we need to consider carefully. Perhaps without realizing it most of the time, we believe that our own experience is somehow the ultimate arbiter of the truth of empirical claims. That this is the case is not in doubt to me, but a short story may help define the issue. A few years ago at this meeting I lis-
tened to a paper on race and mental health, based on our data in Toronto (Tracey-Wortley and Wheaton 1997), and because in fact the findings were counter to what one might find in American cities, suggesting that all of the visible minorities in Toronto have better mental health than the majority white population, disbelief was certainly one of the reactions of the audience. The comment of someone behind me summed up the problem well: “Well, if the stress process model is going to produce findings like this, then what good is it?”

The problem, as I see it, is that we simply think we know better—we have not yet escaped the rule of reason over evidence. Of course, this is a cornerstone problem that hampers many areas of sociology. If we find unusual findings, we tend to reject the finding, often without thinking. If we find what we want, we stop investigating. Both follow from the same problem: assuming we know better.

I wonder whether we stop to think about what is on the computer screen when the findings stare back at us. One thing they stand for is a thousand silent voices speaking together as a coefficient. These are people we know nothing about and may not have access to in our regular social networks for the rest of our lives. We should value and respect this access. If we realize that our own lives amount to a very bad research design—by which I mean it is just my experience with all its uniqueness that I sense, and it is fatally flawed as a route to truth by the roads not taken, that is, by the absence of the important “control group” information I need to reach a better conclusion—and if we put aside “your truth and my truth” notions that obscure the commonality in the origins of suffering in a society, I think we should conclude fairly straightforwardly that experience is not a god, but a very fragile source for leads and nothing more. This is not just an issue of the integrity of research; it is about the importance of knowing what is in fact the case.

I remember back in the 1970s an introductory text in sociology, written by two well-known and respected California sociologists, which dared to discuss criteria for truth. One of the criteria was that “Truth should be beautiful.” No, truth cannot be subservient to our needs to feel good about the world. We ironically fail in our collective mandate to identify sources of misery when we do this. A few years ago, I was using the National Longitudinal Survey of Children to analyze the effect of maternal work patterns on child distress up to the age of 22 (Wheaton 1993). A particularly disturbing finding was that working full-time through the pre-school period had a net positive effect on daughter’s distress up to 18 years later. This is a politically charged finding, and yet I argued that it primarily reflected institutional lag in social change, making the job of parenting particularly difficult at this point of history. This finding raises a difficult question, if it is right: If a result makes us uncomfortable, because it is inconsistent with what we want to believe, should we then suppress it so that we can feel better? If we do, we risk ignoring a problem that is in fact there—presumably the last thing we want to do.

I want to turn to a related example which, for me, has a couple of different lessons. This example involves reporting to you findings from the as of yet unpublished dissertation research of Kim Hall—a recent Ph.D. student of mine at the University of Toronto (Hall 1999). One part of her dissertation was a meta-analysis of studies focusing on the differential effect of marriage on mental health by gender—an issue that has had prominence in public discourse for over a quarter century, and an issue that has been raised in countless and diverse contexts as the driving engine for just as diverse an array of conclusions about marriage and gender. The broad conclusion heard most often, and certainly as manifest in the public consciousness of this issue, is that marriage has fewer benefits for the mental health of women compared to men. Since I assume that sociology should act as an informing and consulting discipline, the rate at which this conclusion is stated is indeed evidence that what we do does have power.

This is an area where intuition runs head on into method. As Hall’s dissertation argues persuasively, we cannot assess this issue by a set of sequential comparisons of given groups. For example, this approach might utilize t-test differences among never marrieds (i.e., unmarried men vs. unmarried women), followed by t-test differences among marrieds (i.e., married men vs. married women) to build an argument for the gender-differential effect of marriage. The problem with these tests is that they do not assess the hypothesis at issue directly; rather, they really only reflect the respective within-role differences,
and thus should not be used to infer a gender difference in mental health gain due to marriage. The hypothesis of a gender-specific effect of marriage requires a single assessment of the male difference for marrieds vs. never marrieds minus the corresponding female difference, in other words, the interaction effect between gender and marital status. Hall also calculates this differential gain coefficient, as I call it, based on 213 independent parameter estimates derived from 78 studies done from the 1930's up to the 1990's. These studies vary widely in their characteristics, from populations sampled to design issues to outcome measures, so Hall also studied the effects of study parameters on findings. The bottom line conclusion is that across all studies, only 14 of the 213 estimates of differential gain are significant, and the average effect size was a very small .016. Thus, there is in fact very little evidence for a differential gain effect benefiting men. Type of mental health measure, disorder vs. scale, positive vs. negative, etc. does not predict substantial differences in this conclusion. However, the year of publication of the study does have an effect, with declining differences over time. Despite this evidence of a direction to the difference, and following the logic of this discussion, this does not in itself imply a difference in earlier studies. One conclusion I take from this is that the intuitive approach to this question has misled us; after all, intuition is only what we know so far.

One particularly important feature of her meta-analysis deserves attention in this context. She reports that studies which target gender-specific gain as their primary hypothesis are also more likely to find evidence in support of gender-specific gain. On the other hand, in studies with a broader mandate, for example to study a range of social factors in mental health, findings about differential gain have been consistently unsupportive of the hypothesis. This could reflect publication bias problems, investigator bias problems, or both.

I believe this example is the tip of the iceberg of a broader problem. There are no doubt many reasons why misleading findings persist, one of which is that they conform to the current Zeitgeist about what is supposed to be. My advice here is simple: We need to be bolder in our work about re-directing the inertia implied by given findings.

Pursuing Ideas

Our area has been a rich source of ideas that indeed define new theoretical perspectives and clarify substantive issues, both within and beyond sociology. Original thinking in particular can place heavy demands on evidence. At times, we set aside (or ignore) concepts or theories because the concepts they invoke are difficult to measure. I see this as the tail wagging the dog. If, for example, we believe that chronic stress is more difficult to measure than life events (as it is), perhaps because it appears to have more subjective or evaluative components, we may choose not to measure chronic stress at all. However, leaving out an essential (and related) piece of the explanation is not really a more conservative approach, since it results in biased estimation of the effects of the concept we believe is measured well. This is so because the unmeasured effect is still lurking in the error term of our regression equations, operating as systematic (not random) error that is correlated with the factors we do measure, thus leaving our results fundamentally ambiguous. I am not advocating the use of bad measures—simply more balanced judgment and thus the inclusion of at least partially successful measures in what we attempt to estimate.

To turn to a specific example of an idea that may be difficult to specify easily in terms of evidence, but which I believe may be important, let us consider a less obvious variant of the idea that social context affects mental health. The general arguments for the direct impact of social context on mental health have recently been articulated in this area by Aneshensel and Sucoff (1996), and I believe that a fundamental part of our collective currency at this point in history, both in our inter- and intra-disciplinary roles, is tied to this idea. After all, it expresses the original issues involved in situating mental health as a part of sociology in the first place. The recent interest in contextual effects is a response to a number of problems, including: (1) the problem of finding and articulating structural effects at the structural level, without implying reductionism; (2) the addition of a layer of social causation that cannot be absorbed nearly as easily by biological alternatives; and (3) the clear separation of individual level manifestations of social position from contextual specification of exposure to a social structure.

Along the way, we should be careful to con-
sider the nexus of social context and social change, as well as the many levels to which the term "context" refers. Someone recently stopped me cold about the promise of this area of research by pointing out that we will probably be too reliant on existing census or aggregate statistics as a source in specifying contextual variation, and that much of what might be important about social context will never be measured unless studies were specifically designed and targeted to address the essential subtleties and varieties of meanings of social context. As a result of this comment, I began to think about what may be missing in our developing collective focus on social context. Like all of us, I presume, I rely in some way on experience for leads, and I have been around long enough to actually be a firsthand observer to some social changes.

There is a layer of social context, of social reality, that may be invisible in the current approaches, which tend to rely on naturally occurring and well-defined social groupings. In articulating what this level of social reality involves, I think back to the work of Erving Goffman and his presentation of interpersonal reality as a realm with its own social structure (1959; 1974). Interpersonal reality is at the borders of our current conceptions of social context. It includes the most proximal and primary groups we are a part of—our families, workgroups, our friends—but it is a wider circle than even the term "networks" encompasses. Basically, this is the reality formed by the aggregation of all our interpersonal experiences in day-to-day life, including not only role enactments but also the casual contacts of day-to-day life, and the observation of others we are privy to by virtue of our social location defined by status, place, and history. Interpersonal reality, I believe, is a hidden social origin of mental health, but its totality is rarely specified or measured in research.

A difficult but interesting question concerns how, and what types, of social changes may be influencing the overall prevalence of mental health problems. Some large-scale epidemiological surveys and some repeated cross-sectional surveys suggest that mental health problems may be increasing. Answers to this question point, quite reasonably, to changing economic circumstances, changing gender roles, changing family structures, or changing rates of crucial stressors over time. Granting the plausibility of these explanations, my sense is that we have to think further about this. What if a significant change in the importance of social context has occurred at the level of interpersonal reality? What would we say this change involves? It is difficult to specify exactly, and I realize that we may all have slightly different takes on this issue (it is, after all, only experience informing us). I have three inter-related possibilities to offer for consideration: (1) an increase in the rate of regulation of daily interaction, (2) the spread of mistrust, and (3) an increase in prevalence of early and easy labeling. In sum, we have evolved, somehow, from an innocent until proven guilty interpersonal model to a guilty until proven innocent interpersonal model.

Regulation occurs at many levels of social reality, and has been extensively analyzed at the macro-system level. A more interesting possibility for me is that the increasing regulation of interpersonal reality has had the net effect of constraining opportunities for the maintenance of mental health at the individual level. Interpersonal reality plays an important function in mental health: Unpredictability, personal expression, and the ability to negotiate "side agreements" that deviate from the larger constraints of social life occur at this level. Regulation of the meaning of interaction, sometimes without regard to the intentions of the actor, is a recent feature of the social climate. My concern is that we cannot standardize the meaning of the subtleties of social interaction, we cannot filter these complexities through pre-coded formalities or legalistics, without risking much about the mental health of the population. Furthermore, I am concerned that a highly regulated approach to social interaction, designed to prevent exposure to that which we are offended by, has the ironic effect of undermining coping resources and thus results in greater population vulnerability to the effects of common stressors.

Before this is automatically coded as an inherently conservative perspective, two things need to be pointed out. First, regulation is obviously a net good, as well as necessity, where the basic protection of rights are involved. Thus, the effects of regulation on collective mental health may be nonlinear, involving a net gain in mental health of the population over earlier increases in the implementation of regulations, but reaching a threshold beyond which collective mental health is compromised. Second, interpersonal reality should
not be assumed to be a direct function of changes in the larger social climate. That is, I am not discussing changes in the rates of law-making or regulation via governmental agencies in general. Rather, the regulation of interpersonal reality may depend more on social change which occurs only at lower social units of analysis, and thus may be expressed typically within organizations, for example, or in the realm of normative change in the informal rules of daily interaction.

Related to the presumed need for greater regulation is the spread of mistrust as a regular feature of daily interaction. Mistrust is an interesting poison. Increases in crime and threats to security, the erosion of authority, the progressive re-definition of private into public life, and a generalized increase in at least the perceived rates of interpersonal betrayal have fed a culture of mistrust. Much of the thinking driving this movement is well-intentioned, as in streetproofing young children. Little is said about the downside of the teaching of constant wariness, of the Hobbesian models of the social world we are socializing. The truth is mistrust is a root source of social disenagement, of hostility and anger, of blame of others, and, ironically, of a kind of immobilization of self. The argument that in fact the world is less trustworthy now is beside the point: Trust works as a basis for mental health even when the recipient doesn't deserve the trust. Remembering both Goffman and the work of Taylor and Brown (1991), we can say that trust is an integral component of the protective illusions of daily life.

A plausible result of these changes is a corresponding increase in the rate of labeling in day-to-day life. I do not refer here to the commonly studied explicit labeling and categorization of persons as, for example, “deviant,” but rather a subtler and more insidious form of labeling that is manifest primarily as rejection, avoidance, or condemnation based on a presumption of knowledge of a person's intentions, attitudes, or values, without actual evidence. The evolution of the interpersonal realm through time is largely described by the notion of an increasing rush to judgment, and an accompanying decline in latitude in behavior. Once the label is applied, the direction of interactions are altered permanently.

If this reasoning has some correspondence to changes that have actually occurred, the issues raised would be admittedly complex to capture with available data. We would need to specifically target the concepts and the level of social reality of concern and find ways to design studies to tap into that contextual reality. I would not suggest interpersonal reality is an important layer of social context if I believed it could not be measured. This is not the place to review possible measures in detail, but I will note two opposite measurement strategies, utilizing the fact that there are both overlapping and unique individual components of interpersonal realities. On the one hand, one can imagine measures derived from aggregated (by place or by time) rates of civil lawsuits, of the rates of regulation of daily interaction in organizations, of the spread of regulatory committees designed to oversee interpersonal claims, or even the aggregation from random samples of averaged trust inclinations. Each of these indicators speaks directly to the issues people face in their day-to-day sensed interpersonal reality. On the other hand, and working from the opposite end, because the individual is at the center of his or her interpersonal reality, we also could derive measures from the reports of individuals in surveys, if we ask the appropriate questions. This might involve measures of daily interaction rituals (typical encounters), and the perceived rules, variability, and content of those rituals. Whether one would want to separate the individual in these reports from the interpersonal environment is an open question, but if so, it may be necessary to use such reports in aggregated groupings as a way of capturing the interpersonal environment as context.

Taking the foregoing speculation more broadly as a signal, we will need to direct the development of studies of social context towards the broader possibilities.

CONCLUDING COMMENTS

It is appropriate to thank the long line of supportive and encouraging influences in my professional life, and especially Len Pearlin for simply being himself. I hope my comments make obvious the fact that sociologists have a unique role to play in the larger effort to track and explain the causes and consequences of mental health problems. My own personal view is that, as an area of research, we need to take a more aggressive stance about the worth of what we offer, and we need to make our work more
difficult to ignore by directly questioning the assumptions or conclusions in existing work. In case it is not obvious, I do not advocate that we need to take this stand within our area as much as across areas and disciplines.

If there is a "latent variable" in this talk, it is the interaction between two themes: structural position vs. analytic power. On the one hand, we look outward to other mental health disciplines from the unique but difficult position of less structural leverage coupled with considerable analytic power. This position implies a responsibility to help guide the study of mental health, but also barriers in the ability to do so. On the other hand, we look inward to the center of our own discipline, from much the same position, knowing that the issues raised by the sociology of mental health are fundamental to the mission and perspective of sociology as a whole. Let us be clear about this much: We should approach the future on the assumption that the worlds beyond our immediate network, in other areas of this discipline, in allied disciplines, in government, and in public discourse, need very much for us to do what we do best, whether they really know it or not.

REFERENCES


Blair Wheaton is Professor of Sociology and Director, Institute for Human Development, Life Course, and Aging, at the University of Toronto. His current research is on the life course consequences of childhood stress, the effects of social context on the relationship between stress and coping, and the forms and conditions of intergenerational influence of parents on children.