Health Care Analysis (2019) 27:289–308
https://doi.org/10.1007/s10728-019-00367-9

ORIGINAL ARTICLE

'Effective' at What? On Effective Intervention in Serious Mental Illness

Susan C.C. Hawthorne[1] · **Anne Williams-Wengerd**[2]

Published online: 20 March 2019

Abstract

The term "effective," on its own, is honorific but vague. Interventions against serious mental illness may be "effective" at goals as diverse as reducing "apparent sadness" or providing housing. Underexamined use of "effective" and other success terms often obfuscates differences and incompatibilities in interventions, degrees of effectiveness, key omissions in effectiveness standards, and values involved in determining what counts as "effective." Yet vague use of such success terms is common in the research, clinical, and policy realms, with consequences that negatively affect the care offered to individuals experiencing serious mental illness. A pragmatist-oriented solution to these problems suggests that when people use success terms, they need to explain and defend the goals and supporting values embedded in the terms, asking and answering the questions, "Effective *at what*? *For whom*? *How* effective? And *why that goal*?" Practical and epistemic standards for effectiveness will likely remain plural for good reasons, but each standard should be well explained and well justified.

Keywords Serious mental illness · Treatment · Effectiveness · Psychology · Psychiatry · Clinical social work

✉ Susan C.C. Hawthorne
schawthorne@stkate.edu

[1] Philosophy Department, St. Catherine University, 2004 Randolph Ave, St. Paul, MN 55105, USA

[2] Psychology Department, St. Catherine University, 2004 Randolph Ave, St. Paul, MN 55105, USA

It's easy to agree that it is a good thing to be effective—or, in a more complete phrase, to be "effective at achieving goal X"—without carefully considering what X is and why one would want to achieve it. People involved in the scientific, clinical, and political contexts of intervention in serious mental illness (SMI)[1] commonly use "effective" or similar success terms with such generality. By doing so, they can speak both within and across disciplines, where key negotiations about what to do, why, and how take place.

Yet without detailed specification, use of the term "effective" achieves the rhetorical purpose of making claims with which it is difficult to disagree, but it introduces obfuscations. For example, the American Psychological Association's 2006 position paper on evidence-based practice (EBP) states in a sentence typical of the document, "Meta-analytic investigations since the 1970s have shown that most therapeutic practices in widespread clinical use are generally effective for treating a range of problems" [2, p. 274]. The authors do not specify the type of problem, conditions needed for effectiveness, degree of effectiveness, or measurements used.[2] Presumably, they considered such issues to back their claim, but the rationale is opaque to readers, whose criteria for effectiveness might differ from those applied.

Through its blandness and ubiquity, such vague use of success terms tends to escape attention,[3] while simultaneously hiding details about the goals achieved, the reasons these goals are worth achieving, and the worthiness of goals not attempted or achieved. Thus, such use introduces conceptual puzzles and concerning consequences that warrant analysis. One conceptual puzzle centers on the inconsistent, incomplete, and sometimes incompatible standards used to benchmark effectiveness across fields involved in SMI intervention. For example, evidence-based medicine (EBM) standards for effectiveness differ from those espoused by individualized medicine; successful outcomes in genomics research little resemble those in services

[1] We use the definition of serious mental illness proposed by the National Institute of Mental Health (NIMH): "a mental, behavioral, or emotional disorder resulting in serious functional impairment, which substantially interferes with or limits one or more major life activities" [46].

[2] The document does sketch a range of questions that need addressing in research, such as "the efficacy and effectiveness of psychological treatments with children and youths at different developmental stages," [2, p. 275]. It also notes the variability of interventions—for example, the report notes that "… psychological services are most likely to be effective when they are responsive to the patient's specific problems, strengths, personality, sociocultural context, and preferences" (p. 278). Even with such refinements, however, the criteria for effectiveness can be specified in multiple ways.

[3] Perhaps this blandness explains why the use of general success terms like "effectiveness" has not been much examined in the literature we surveyed. Sandra Tanenbaum's work is a notable exception. Within the context of critique of evidence-based practice, Tanenbaum presents well the basic insight that the definition of "effective" is both unclear and contentious [62]. She points out that "effectiveness" is defined in objective terms in EBP, and that these measures to do not reflect treatment outcomes pursued by many. She also raises the critical issue of who is defining effectiveness, and she points out the potential coerciveness of implementing treatment or policy with "effectiveness" defined in certain ways. Our discussion is broader than Tanenbaum's, notably in pointing out that the need to define effectiveness is critical for all involved in SMI intervention, not just those doing EBP. In addition, while Tanenbaum points out that "what works" has moral connotations, she doesn't enlarge on this or explain the moral connotations of other approaches, nor does she demand specification of goals. Cartwright and Stegenga are another exception; they consider the term in discussing criteria and methods for translating scientific efficacy to policy effectiveness [10].

research; and a policy goal of safe housing might not align with a goal of low cost. Vacuous use of success terms also helps hide how little is known, by any standard, about effective intervention against SMI. A second conceptual puzzle centers on identifying the diverse interests and other values held within the varying standards by which interventions are said to be effective. The poor articulation of effectiveness standards, and of the values embedded in them, feeds conflicts and misdirections that contribute to poor understanding and resolution of scientific, clinical, and public disagreements.[4]

Such consequences impact the lives of people experiencing SMI. Part of a complex of influences—others include compassion, money, politics, power, stigma, knowledge, and ignorance—they together have shaped a fragmented, often inaccessible, and sometimes inhumane US mental health system. (We write with that overarching concern in mind, but we expect that much of our analysis is relevant in other cultures.) Despite that specific concern, our analysis does not argue for or against any particular criteria for effectiveness, or against use of success terms. Instead, we suggest that to avoid conceptual traps and their practical outcomes, scientists, mental health care providers, mental health policy regulators, service users, and others who wish for effective intervention against SMI need to confront vacuous usage, asking, "Effective *at what*? *For whom*? *How* effective? And *why that goal*?"[5]

We do not wish to overstate: the people involved with SMI intervention do recognize the upside of greater specificity in some ways. Intervention researchers, for example, distinguish between "efficacy" research (effective within the confines of the study's controls) and "effectiveness" research (effective in clinically realistic settings). Similarly, statistical significance is commonly distinguished from clinical significance, and this distinction has opened a debate about how to conceptualize and measure clinical significance [22, 30, 37, 58]. Some debates over the usefulness of evidence-based medicine ask whether it's more important to know what's effective for populations (an average result), or whether the goal should be knowledge of what's effective for varying individuals [35]. More generally, scientists attend to specifying criteria for effectiveness whenever they operationalize outcomes; clinicians, in effect, set parameters for effectiveness each time they set a therapeutic goal; and policy makers debate the relative merits of various choices.

Nevertheless, further clarification of meaning can highlight both bases of disagreement and shared goals around "effective" intervention against SMI. We thus offer our criticisms and suggestions hoping to speak to those who share our concern about troubled systems of mental health care in the US and elsewhere, in hope that the work can ultimately benefit those experiencing SMI. Our paper first develops, in the section that follows, a more complete description of the vague use of success terms in the clinical, research, and policy realms, showing ways in which

[4] A third conceptual issue concerns the complexities of translating scientifically demonstrated efficacy or clinically demonstrated need to effective policy; on this issue we offer a pragmatist-oriented direction, but details are beyond the scope of this paper.

[5] With respect to psychological research, Paul independently developed a similar statement of the problem in 1967 [51].

the underlying standards for success are incomplete, misaligned, or incompatible. The "Consequences of Under-Considered Usage" section explores in more detail how values underlie expressed goals in each venue, and further draws out negative impacts of underexplicated use of success terms. We then offer, in the concluding section, pragmatist-oriented recommendations for improvement—basically, that when people use the word "effective" or similar success terms, they explain and defend the goals and supporting values embedded in the terms. Practical and epistemic standards for effectiveness will likely remain plural for good reasons, but each standard should be well explained and well justified.

What is 'Effective'? Divergent Understandings

Modes of intervention against SMI include client care, research to guide improvements in intervention, and policy that guides both. These modes interact, of course. This section explores the diverse effectiveness standards [62] in the realms of client care, research, and policy, along with some of the cross-realm interactions.[6]

Diverse Goals of Care Providers, Clients, and Non-clients

On average, clinical practitioners of psychiatry, psychology, social work, and related fields active in SMI intervention emphasize different goals, according to the expertise and ethos of the field. Psychiatrists most commonly have a biological orientation, emphasizing symptom reduction; clinical social workers adopt either mixed models of intervention or a psychosocial model favoring recovery goals, such as meaning making [9, 52]. Most psychologists focus on the individual experience, basing endpoints primarily on self-report.

However, different schools of practice within fields emphasize different goals. While biological psychiatry targets specific neurobiological malfunctions, phenomenological psychiatry targets subjective experience [24], and recovery advocates within psychiatry argue that psychosocial supports such as employment and recreational opportunities are medically necessary [29]. Schools of psychology are at least as diverse: Cognitive behavioral therapies tend to emphasize rationality and short-term care [6]; psychodynamic therapy aims at developing capacities for flexible engagement and creative exploration, emphasizing client diversity; existential therapy's therapeutic goal is to enact freedom, authenticity, angst, and awe, to help clients "better face the conditions of human life [6, p. 1079]." Within social work, too, approaches are debated—in particular, whether to accept an EBP paradigm over a psychosocial model [24].

This diversity of aims sometimes avoids sharp conceptual or practical divisions, given overlapping goals, mixed models, and teamwork. Often, however, the goals of

[6] Some "non-clients" are people experiencing SMI who cannot access care. Others are non-clients because they prefer to avoid intervention, tend to their own care, or work with peer interventions; their perspectives are represented in this section and elsewhere.

the different schools are not only diverse, they are incompatible. For example, many clinically and personally significant incompatibilities arise between those providers and clients who adopt a recovery model and those who work under a biomedical model. The term "recovery" "…refers to a person's right and ability to live a safe, dignified, and meaningful life in the community of his or her choice despite continuing disability associated with illness" [14]. It is "…more of an attitude, a way of life, a feeling, a vision, or an experience than a return to health or any kind of clinical outcome per se" (Deegan 1998, 1996a, cited in [15]). Effective intervention or "self work" promotes the ability of affected people to experience recovery so defined; the experience is a process at least as much as an endpoint.[7] The biomedical model of mental illness, in contrast, considers the core difficulty of SMI to be symptoms caused by a biological dysfunction. This biological dysfunction, a feature of the individual—is (in principle) identifiable, causally explicable via scientific methods, and amenable to medical treatment [27, 55]. Under this model, effective interventions alleviate symptoms and, ideally, also manage or repair the biological dysfunction; that is, they achieve a specific version of "return to health" and specific clinical outcomes. Thus, a central incompatibility is that recovery model supporters argue for individuals' moving forward with life goals while experiencing symptoms, whereas biological model proponents see the process of improvement in stages: First decrease symptoms, then tackle other life goals [13]. To take a more specific example, under a biological model individuals treated for schizophrenia will almost always be prescribed an antipsychotic to help control the positive symptoms of schizophrenia, such as delusional thinking or hearing voices. For some individuals, antipsychotic use and recovery are complementary because symptom relief helps them achieve their recovery goals. For others, however, antipsychotic drugs interfere with recovery goals: The sedative effects may interfere with cognition [4]; side effects of weight gain, lethargy, or type 2 diabetes may decrease physical health. The drugs may also decrease creativity, cause emotional blunting, or diminish sex drive. Any of these factors may interfere with participation in activities that make life meaningful.

Rather than being incompatible, goals may be misdirected. On average, clients rank accessibility of social resources above goals related to symptom change. For example, despite being aware of EBP and clinical goals, they emphasize practical goals such as crisis stabilization, jobs, housing, education, friends, cash, and Social Security funds [11, 21, 63]. Destigmatization [63] and greater acceptance of diversity [43] are other aims reflected in surveys of people experiencing SMI. When the health care system emphasizes symptoms rather than life experience, or when resources are unavailable for needed assistance toward client-desired ends, goals are misdirected from clients' perspectives.

[7] The term "effectiveness" implies an endpoint, while "process" and "relationship" do not necessarily. Nevertheless, processes and relationships can be better or worse (i.e., more effective) at achieving other goals, and more or less satisfying in the present (hence, more effective at an overall goal of a life well lived). Attempts to assess the quality of processes and relationships are therefore common in research and practice.

Some disagreements confronted in clinical choices seem to concern means of intervention, rather than goals. For example, clients and providers might agree that stabilized symptoms, continuity of care, and fewer hospitalizations are goals worth achieving, yet differ over civil commitment to assisted outpatient treatment (AOT). Many professionals argue that for individuals meeting the criteria, mandated AOT best achieves these goals; opponents resist the entailed deprivations of freedom [31, 45, 49, 59, 61]. In a second example, the need for "treatment" is contested: Groups like the Hearing Voices Network [47] argue that professional involvement may be detrimental, and they advocate for peer support or self-care.[8] While not denying that the goals/means distinction is useful in some settings, our framing interprets the disagreement as concerning additional goals. For example, AOT proponents might propose a goal of consistent intervention; treatment opponents might wish to maximize independence. In this framing, aspects of effectiveness—and disagreements over them—come to attention that might not otherwise.

These examples demonstrate that the diversity in aims across fields and schools of thought is not simply difference. Instead, it entails mismatches in goals and misdirections of care with potentially negative consequences for those experiencing SMI. These consequences are further explored under "Consequences of Under-Considered Usage" section below; recommendations for change follow in the concluding section.

Research Endpoints, Research Aims

While the goals of scientific research on SMI are equally divergent, the plurality of SMI intervention research in the US is done under a biological model, with much emphasis on pharmaceutical interventions.[9] Under these influences, research questions, operationalizations, and resulting recommendations for intervention tend to feature aspects of SMI plausibly addressed by medication. That is, the bulk of the research investigates symptom reduction. In doing so, it leaves out goals emphasized by recovery, phenomenological, or psychosocial models. Similarly, prevention has not been a focus.

Reliance on the DSM classification system for identifying research subjects is another trend that misaligns much research with clinical and personal goals. We will not review all the purported flaws of the DSM here [53]. One key concern, however, is that the DSM's descriptive and pragmatic classification system does not reflect distinct physiological or psychological differences among people. The lack of alignment means that people in any DSM diagnostic group are heterogeneous in ways

[8] Additionally, a few bioethicists argue that effective *treatment* goals are not appropriate for all clients with SMI. For example, Wijsbek argues that euthanasia may be appropriate for intractably depressed patients [65], and Draper suggests that treatment refusal, rather than forced feeding [18], should be allowed for some patients who suffer from severe and prolonged anorexia nervosa. It is not as clear that providers endorse such positions.

[9] Drug companies have in recent years decreased their spending on psychiatric drug development. This is not necessarily a good thing, given that improved medications, with fewer side effects, could be useful.

that are not understood, muddling research based on the category. Further, attempts to control for heterogeneity—for example, through the typical practice of excluding people with comorbidities from research [32]—leads to non-representative samples and poor external validity (that is, poor modeling of real-world problems), given that nearly half of all persons experiencing mental illness meet the criteria for more than one DSM-based diagnosis. Similarly, the predominantly quantitative research base usually determines average or typical responses to interventions. In a field where variability is the norm, the averages misrepresent a large number of responses to intervention. In addition, the DSM-5 has turned further toward an implicitly neurobiological understanding of mental illness by removing the multi-axial system that earlier editions used to understand individuals' functioning in psychological, social, and cultural context. In doing this, the DSM-5—and the research predicated on its categories—downplays factors that contribute substantially to understanding the different responses of treatment groups—factors that many clients and providers say matter the most. Thus, whether the goal is biological or psychological accuracy, or intervention targeting endpoints deemed significant by those affected, using DSM-5 categories introduces misalignments between the research and other goals. Even the rare[10] evidence considered strong by EBM standards [39, 44] often says little about effectiveness defined according to goals many would endorse, such as obtaining housing, improving relationships, or obtaining meaningful employment.

Recognizing these difficulties with using the DSM categories in research, the National Institute of Mental Health (NIMH) is gradually moving away from funding research based on them. However, its budget prioritizes its own Research Domain Criteria, which focus on understanding the physiological mechanisms of mental illness [54]. The NIMH budget thus also places low priority on research tied to recovery-model or client priorities, such as effectiveness of social service or community-based interventions, physical health, and prevention efforts [20, 26, 56]: only 10% of the NIMH research budget is devoted to such research [26].

Individual intervention studies present a similar concern—the poverty of operationalizations in representing clinically or personally significant goals. Ironically, this concern most often arises for (arguably) scientifically justifiable or unavoidable methodological reasons, including the ethical and practical realities of research, the

[10] The knowledge base on interventions for serious mental illness is weak even taken according to commonly accepted "rules"; that is, accepting DSM categories and the standards of evidence-based medicine. For example, the 2005-2015 Cochrane database entries for bipolar disorder and schizophrenia find "strong" evidence that antipsychotic medications relieve positive symptoms of these disorders in the short term. But that is the only strong evidence found among hundreds of articles that review dozens of medication types, psychosocial interventions, and endpoints.[Unpublished analysis by one of the authors, SCH] Morlino et al. [44] undertook a study of systematic reviews of interventions for schizophrenia in more databases, including the Ovid and Pubmed databases as well as Cochrane. They determined that only 25% of the reviews produced definitive conclusions. In addition to these concerns, Leucht et al. also point out that the entire Cochrane database (comprising meta-analyses including 13,542 articles, by their count) concerns only "first pass" efficacy—that is, the articles consider the effectiveness of initial interventions, but don't provide guidance on the clinically crucial question of what to do if the approach fails [39]. Turning to the literature, the group found only 10 articles that addressed switching antipsychotics, and found that none of these gave conclusive information.

specialization of scientific fields, the difficulty of operationalizing thick concepts, and the choice to use qualitative or quantitative methodologies.

The practical and ethical realities of research frequently require methodological compromises. Practical limits on money and time mean that any individual study can only take on so much; endpoints may be left out not because researchers consider them insignificant, but because resources to track them are not available. Ethical concerns also limit viable options. To avoid coercing vulnerable subjects, for example, researchers might study a short-term endpoint when long-term results would be more clinically significant.

Researchers' decisions about what to investigate and how are motivated not only by clinical need, but also by factors specific to their fields. These include the types of questions characteristic of their specialty, intellectual and tangible resources, methodological choices, and reward structures [34]. For these and similar reasons, pharmacologists and neurobiologists do not study housing; social scientists do not study neurotransmitters or use fMRIs. In principle, the various fields can contribute pieces to understanding effective SMI intervention—but gaps and misalignments are common.

Aware that individual measures provide a limited view of effectiveness, researchers typically try to get a fuller picture by grouping several measures within a study. While often helpful, the multiple measures may still be of dubious value, resulting in dubious claims of effectiveness. For example, the pharmaceutical company AstraZeneca funded two studies of the extended-release version of their antipsychotic medication Seroquel (quetiapine) that investigated its use as adjunctive therapy for major depressive disorder [5, 19]. The studies employed a multi-center, randomized, double-blind, placebo-controlled design, and sought to establish effectiveness at two dosage levels, 150 mg/day and 300 mg/day. The researchers used several rating scales, completed by people with varying perspectives: Patients completed the Hamilton Rating Scale: Depression (HAM-D) and the Quality of Life Enjoyment and Satisfaction Questionnaire (Q-LES-Q), the primary psychiatric care provider filled out the Clinical Global Impression Scale (CGI-S), and a third-party clinician completed the Montgomery-Asberg Depression Rating Scale (MADRS).[11] There was no change in the patient-completed Q-LES-Q in either study. In one study only [19], researchers observed change in the HAM-D and CGI-S; and the third-party completed MADRS scores changed significantly at 300 mg in one study [19], and at both dosages in the other [5]. Despite the lack of demonstrated effects on quality of life and limited evidence of change notable by the treating clinician, the authors concluded that, "Quetiapine XR 300 mg/d was effective against a range of depressive

[11] The Montgomery-Asberg Depression Rating Scale (MADRS) is scored on a 60-point scale. The items examined include apparent sadness, reported sadness, inner tension, reduced sleep, reduced appetite, concentration difficulties, lassitude, inability to feel, pessimistic thoughts, and suicidal thoughts. For example, the "apparent sadness" section assesses **"**despondency, gloom, and despair (more than just ordinary transient low spirits) reflected in speech, posture, facial expression, and posture." The rater is instructed to "rate by depth and inability to brighten up" with the following scale: 0 points for "No sadness"; 2 points for "Looks dispirited but does brighten up without difficulty"; 4 points for "Appears sad and unhappy most of the time"; and 6 points for "Looks miserable all the time. Extremely despondent."

symptoms at week 6 in this patient population, with onset of efficacy seen at week 1 [19, p. 930]," and that "Adjunctive quetiapine XR (150 mg/day and 300 mg/day) was effective in patients with MDD with patients who had shown an inadequate response to antidepressant treatment [5, p. 540]." Subsequently, the FDA approved quetiapine as an adjunctive therapy for major depressive disorder despite the known serious side effects of antipsychotic medications. The scientists' claims and the FDA's approval are not atypical: Discounting the clinical and personal importance of deleterious side effects is common.[12]

More basically, however, the thick concepts that are clinically and personally meaningful—e.g., relationship, connection, or quality of life—are complex and not easily measured. Inevitably, operationalization incompletely represents the concepts. Even biomedical notions such as "symptom reduction" raise complex issues around deciding what counts as a "symptom" and which symptoms matter most. The concepts used in recovery models and public health approaches are broader and more contentious: For recovery models, what is "meaning," and how does one measure it? What is "public safety" or "community engagement," and how should they be quantified? One *can* operationalize these [57]—it's done all the time, directing attention to parameters the scientists involved can study and that they consider important, given their resources and interests. But no reasonably practical operationalization can capture the richness and diversity of views on these thick concepts. And that failure leaves the content vague and incomplete when interventions to restore meaning-making, public safety, community engagement, and the like are said to be "effective."

Both because it is difficult to operationalize thick concepts, and because many aspects of doing so are not engrained in practice, research done under non-pharma, non-DSM, or non-biomedical models may do little better at specifying effectiveness standards. For example, under a recovery model, Yanos et al. [66] aimed to study the "successes and challenges" of housing in fostering "community integration" for those experiencing SMI. Researchers asked participants about their subjective responses to housing, and elicited comments on feelings of safety, loneliness, and fitting in. Yet, as the researchers themselves point out, their assessment of "community engagement" did not determine whether participants were taking an active role in their communities—arguably a central criterion by which one could judge successful community integration. More broadly, researchers using a recovery model can choose among multiple questionnaires designed to operationalize "recovery," selecting instruments that measure variables of interest with respect to the problem they are studying. However, none of the questionnaires captures the myriad interpretations possible for the term, so none succeeds in uncontentiously establishing criteria for "effective" recovery [23].

Qualitative research often attempts to avoid the "problem" of multiple interpretations by embracing the multiplicity. For example, qualitative studies of intervention often allow individuals to define effectiveness for themselves. Larsen-Barr's analysis of the subjective experience of taking and quitting antipsychotics is a case in point,

[12] Thank you to an anonymous reviewer for making this point.

revealing wide variations in success and insight into a range of reasons people find the drugs helpful or hellish [38]. However, the generally recognized limitations of qualitative research[13] preclude its being an overall solution to the problem of determining intervention effectiveness—quantitative methods are needed, too.

If individual studies can only present a fragmented and debatable view of effectiveness, comparison between studies will often be an "apples and oranges" affair. For example, Padgett et al. [50] favorably compared "substance use outcomes" of "Housing First" to "Treatment First" approaches to SMI intervention. Housing First approaches provide stable housing, then offer SMI and substance abuse treatment; Treatment First requires mental stability and sobriety before permanent housing is offered. Padgett et al. considered reduction in substance use to moderate or occasional use a benefit of Housing First approaches. However, because only abstinence counts as success in a Treatment First model, this Housing First benefit is irrelevant within the Treatment First framework: That is, "effective" Housing First and Treatment First results are not directly comparable. Similarly, much biomedical research assesses short- to medium-term results (weeks to 6 months). Goals within the short-term study are often irrelevant to long-term recovery goals—such measures as drop-out rates, leaving the study early, cost-effectiveness, and narrow symptom measures correlate unpredictably with long-term recovery goals. Andresen et al. [3] assessed this point systematically. They explain, for example, that symptom severity does not correlate with having hope—presence of hope being a significant component of the recovery model. According to Andresen et al., individuals can have severe psychotic symptoms, yet experience a strong sense of hope about their life or vice versa. While optimists might believe that short-term effectiveness can be parlayed into the long term, the realities are not so simple.

Overall, then, existing research presents many vague, misaligned, incomplete, or incompatible answers to the questions, "Effective *at what*? *For whom*? *How* effective? And *why that goal*?" Yet, through banal and slippery uses of success terms like "effectiveness," the extent of the incompatibilities, missing information, and their consequences escapes detection.

Policy Disagreements

Institutional, legal, and legislative policy decisions touch all aspects of the SMI intervention, including research priorities; public and private clinics; public and private insurance; community care and outreach; housing; and the education, legal, and prison systems. By establishing laws or guidelines and by allocating resources, policy decisions related to SMI interventions set parameters for both public and private research and for clinical and social interventions; in the case of private institutions and families, the effect may be to work around regulations or fill in gaps in resource allocation.

[13] Relative to high-quality quantitative research, qualitative research has less control and generalizability. Both qualitative and quantitative research embed biases; interpretive methodologies make qualitative research especially prone to observer biases.

SMI-related policy decisions are complex and contentious, pitting varying conceptions effectiveness and appropriate goals against one another in a context of myriad interactions among systems and individuals, and limited resources and information. Drake et al. [17] provide a US-focused example that illustrates some of these difficulties. They argue that the majority of people experiencing SMI want to work, and that achieving the goal of steady work also effects additional successes, such as increased self-esteem, quality of life, and reduction in mental health services use. Individual placement and support services have been shown to greatly boost employment percentages among adults experiencing SMI. However, it is not as clear that individual placement and support services achieve other goals, such as cost effectiveness or, as an early intervention, altering the course of mental illness. Decisions about who has access to such services, the extent to which these individuals can also receive disability payments, the amount of money available in the Social Security system, the specific protocols for training and support, and many other factors need to be considered in establishing policy. In an equally complex example, many favor both public safety and engaging those experiencing SMI in their communities of choice. But these goals may clash: Policies that favor maximizing public safety may isolate more people experiencing SMI from mainstream society than policies that favor maximizing community engagement [25].

In the common case in which policymakers set goals according to a narrow set of effectiveness standards, significant possibilities based on other standards go underexplored or underexplained. The resulting gaps have negative repercussions discussed further below. For now, suffice it to say that policy makers evaluating a claim that a given policy is or will be effective need to know at what it is effective, for whom, under what conditions, and whether the goals it satisfies are justified.[14]

Consequences of Under-Considered Usage

Overall, then, the diverse people and fields involved in SMI intervention judge effectiveness by standards that are typically unclear, incomplete, inconsistent, or incompatible. The varying standards used reflect the interests, desires, and ideals (the values) held by the millions involved—particularly by those with relative power. The resulting conceptual and rhetorical omissions and commissions have practical consequences that are frequently negative for people experiencing SMI, providers, the knowledge base, and policy. We further explore the involved values and practical consequences in this section.

[14] Establishing ways to evaluate how scientific efficacy translates to clinical or policy effectiveness is a very active area of discussion in the philosophies of science and medicine [10, 16]. The details are beyond the scope of this article.

Research

As we have seen, concerns with the overall research base include the incompatibilities of some effectiveness standards among schools of thought and practice on SMI; limited attention to goals relevant to recovery, phenomenological, or psychodynamic models in the predominant research base; risks of basing research on DSM diagnostic categories; dominance of quantitative over qualitative research suggesting underrepresentation of variation; and the difficulties and failures of representing clinically or personally meaningful goals in measurable form. We have already discussed the practical consequence that the multiplicity of effectiveness standards makes it difficult to compare studies and to apply research results to individual clients.

The consequences go further, however. Given that research endpoints often reflect the interests of commercial or academic sponsors or schools of thought [36] that are not attuned to the needs clients and providers express, some significant portion of the research is poorly targeted to expressed needs. Interests in profit are an obvious example. Another is that the DSM categories that underlie much research embed debatable values by designating certain behaviors or experiences as "symptoms," and thus as negative; these values may not be shared by people experiencing SMI. Poor targeting limits research relevance. It also leaves significant areas of ignorance. For example, people who have schizophrenia live, on average, 25 years fewer than those who do not have the disorder. Lifestyle (smoking, exercise, weight control) might well explain part of the variance, yet there is very limited research on lifestyle interventions [7, 33].[15] Finally, the term "effective" can be used to mislead. The decades-long overstatement of antidepressant effectiveness is an example [28].[16]

Thus, while it is true that weakness or absence of evidence for effectiveness does not prove that interventions are ineffective, and while it is also true that many providers and clients claim significant benefits from current approaches, the considerations just discussed seriously weaken or display weaknesses in the knowledge base relevant to SMI (see also footnote 10). Downstream, the effects on the research affect providers, clients, and policy: Providers are left without adequate information to do their work, clients are ill-served by weak or poorly targeted empirical studies, and policy makers lack the empirical evidence needed to craft "effective" health care policy and resource allocation.

Clients and Providers

As we have seen, providers conceptualize goals and interventions in multiple ways, following their orientation to particular schools of thought or practice. Such orientation is typically also to a set of values. Proponents of the recovery model typically

[15] Given the likely connection of these "lifestyle" variations to mental illness, drug side effects, or both, the term "lifestyle," suggesting full agency and responsibility, is likely misleading here. Still, well-targeted interventions might enhance quality or duration of life for people experiencing schizophrenia.

[16] According to a recent meta-analysis, effect sizes are mostly modest and acceptability is lower relative to placebo [12].

value social connectedness and inclusive attitudes towards neurodiverse people. Those adopting the biomedical model of SMI inherit the core values of medicine—survival, security, and flourishing [40]—which medical practice works to secure at the level of individuals, rather than societies. Public health professionals often hold collective values such as solidarity ahead of liberal and individualistic values [1]. People who experience SMI may share any of these values or may emphasize others. Members of the antipsychiatry movement, for example, value freedom over medical intervention, particularly over mandated treatment. More generally, values of individuals who experience SMI can be inferred from their stated preferences, which tend to emphasize practical goals.[17] In addition, providers and clients are focused or constrained by others' goals and priorities, such as insurance company, hospital, and government concerns with cost-effectiveness.

Given these individual and institutional wants and needs, the wide range of interventions is not surprising—there are many varieties of mental suffering [64], and many ways to address it may be effective. But the varied interventions are not merely options on an enticing menu. First, the contrasts discussed above, such as Housing First versus Treatment First models and meaning-oriented versus symptom-defined endpoints, show that some interventions and effectiveness standards accord with some values, and some with others. Perhaps all involved would describe the overall goal of SMI intervention as enabling a good life—but what's a "good life"? Is it symptom-free? Must it have meaning to the individual? Is participation in society necessary? Second, it is not clear at present for which populations or individuals each form of intervention is most effective (by relevant standards): Adequate research to understand response variations simply has not been done. Finally, the range of options is often unavailable: Individuals may be forced to "choose" an approach incompatible with their values or response profiles.

In criticizing the inapplicability of current research to some individuals, we acknowledge that research has limits that are difficult to overcome. Well-designed trials provide important information for practice, but research is not therapy: Therapy is typically individualized, while research protocols are typically not. It is unfair to demand that controlled studies show effectiveness as the diversity of individuals experience it, and equally unfair to demand that qualitative research exhaust the definitions of thick concepts. Nevertheless, it remains crucial to recognize that the limitations of research—quantitative and qualitative—narrow the perspectives they present on intervention effectiveness.

Policy

As with client care and research, unshared values frequently drive disagreements among and between policy makers and constituents over what policies are

[17] Debates over the rationality of goals expressed by individuals with some forms of SMI, or by individuals at some stages of illness, introduce complex issues involving judgments of goal rationality and whether goals need to be rational. We will not debate that issue here, but we point out that it applies to a minority of people with SMI.

considered effective and over the effectiveness standards by which policy will be judged. For example, many insurers, administrators, and legislators highly value cost control, making "cost-effectiveness" rather than "effectiveness" simpliciter their overall goal. This prioritization, however, is at least partially inconsistent[18] with that of people who argue for resource-intense interventions such as providing individual counseling or supportive housing which, in their view, better support individuals' unique goals and quality of life. Other attitudes and priorities prevalent in US society generally also affect goal-setting: Examples include widespread stigma against people experiencing SMI and US society's embrace of individualism and personal responsibility [48]. Divergence in values similarly drives debates over civil commitment, regulation of drug company profit, and publicly-funded research agendas.

Thus, at the policy level, the consequences of underexamining effectiveness standards parallel those that affect direct care. Multiple approaches to SMI intervention compete, and evidence of effectiveness cannot be well used to craft policy if "effectiveness" standards are poorly, incompletely, irrelevantly, or incompatibly determined. The point here is not to prejudge policy, or to assume a one-size-fits-all approach. However, in the US, the negative practical results of underexamined effectiveness standards (along with other driving forces such as stigma) include mental health care deserts, mass imprisonments of mentally ill individuals, neglect that often ends in homelessness, early death, inconsistent access to needed medication, overreliance on medication, and sharply limited social support. That excellent care is available to some testifies to competing values at work—compassion and respect for human dignity among them. The contrast itself suggests that without deep discussion of what the goals are, why they should be supported, and how well they can be reached, claims of effectiveness are gratuitous and can be used to distract from deeper issues.

Together

Thus, conceptual disagreements over effectiveness standards and the values that underlie them exist across the many realms and individuals involved in SMI intervention. These disagreements—and the failure to work through them—hamper the development of evidence and of clinical and social care. Despite efforts in all realms, the practical effects of these issues on client care, research, and policy are among the reasons that many with SMI experience inadequate, unwanted, or misdirected interventions.

[18] We say "at least partially inconsistent" because complications of up-front versus net costs, among other factors, make costs difficult to determine.

Clarity and Diversity

If our arguments about vague and incomplete uses of success terms are well-founded, the basic solution is straightforward: Be sure to communicate and understand success criteria and their rationale. In practice, of course, the solution is more complex. Here are a few steps:

First, when investigating SMI interventions, setting treatment goals, or crafting policy, recognize that judgments of "effectiveness" or similar success terms involve both factual and values-based claims. The facts represent the observable consequences of any intervention. Does a symptom lessen when taking a medication? Does alcohol intake decrease when housing is provided? For most questions asked about SMI, the facts are still uncertain and require more investigation. But equally important, the choice of which questions to ask, which consequences to observe, and which to set aside embeds positive or negative valuation of the endpoints involved, and the values underlying those goals vary greatly.

In the second step, identify *which* facts and values are driving the choice of operationalization or judgment of effectiveness. Answering the questions, "Effective *at what*? *For whom*? *How* effective? And *why that goal*?" helps determine which facts and values are in play, and it directs attention to potential impacts of the choices or judgments. Determining "at what, for whom, and how effective" helps illuminate gaps between research findings and their clinical significance; greater recognition of the limitations in scope could aid clinical translation by helping to pinpoint who falls within and outside that scope. Answering "Why that goal?" focuses attention on values, and it helps clarify whose interests are represented in the effectiveness operationalizations and goals [62].[19,20] Recognizing whose interests are included and excluded can help all involved focus on endpoints that matter to people experiencing SMI. In addition, answering the "why?" question draws attention to whether effectiveness is even the right framing [62]. "Effectiveness" tends to focus attention on endpoints, implying the desirability of intervention. That focus also potentially discounts process and progress [6], potentially undervaluing longitudinal practices such as routine outcome monitoring [8] or developing therapeutic relationships.[21]

Third, make the answers to these questions public, so they can be debated [41, 42]. Those reviewing research proposals or completed research on SMI intervention, providing direct care to people experiencing SMI, undergoing SMI intervention, or developing mental health policy can demand answers to the questions about effectiveness and justification for the answers.

[19] Tanenbaum makes a similar point in the context of EBP "…whenever EBP researchers state their outcome measures, [T]hey should say why they have chosen those measures and whose interests they represent (p. 171)." [62] Our point is that it is not just EBP researchers that need to do this, but all researchers, and providers and policy-makers as well.

[20] Other questions might also point to values. For example, asking "Effective at what?" and "For whom?" jointly might point toward effectiveness standards that benefit caregivers or investors more than patients.

[21] See also Footnote 7.

These steps are not completely foreign: People establishing operationalizations or implementation already think this way to some extent. Typically, however, their process does not bring values into perspective, or require the degree of specificity and defense suggested here. The suggested extra measures raise the bar in systems where specific effectiveness standards and certain goals have a long history. The steps are not onerous, however, and can readily be part of proposing, writing up, or reviewing research; embarking on an intervention as provider or client; or choosing a new policy.

Importantly, examining values along with facts is compatible with the "untidy pluralism" [60] of knowledge-building practices (basic science, consensus conferences, EBM, narrative medicine, etc.) in medicine and other professions involved in SMI. It is compatible as well with pluralism in SMI interventions—biomedical, biopsychosocial, psychological, personal, sociocultural, and others. Finally, it is compatible with pluralism in research. Indeed, although monism is another way to avoid incompatible goals, values, and endpoints in research and practice, we think monism is the wrong approach. First, the central problem we have identified is *not* that there are multiple hypotheses, questions, goals, and endpoints or multiple types of intervention; rather, it is that the rationales for many of these are poorly articulated, planned, debated, or understood. Clarification would help to garner the key advantage of pluralism, which is that multiple methodologies and questions open doors to potentially useful interventions, models, and ideas that monism would not. Recommending clarification rather than monism also recognizes overlapping goals among various models of intervention. For example, the endpoint of "time to rehospitalization" might be relevant in both biomedical and recovery-oriented research or practice. For the biomedically-oriented scientist or provider, "time to rehospitalization" is significant because it operationalizes a waxing of symptoms; to a recovery-oriented scientist or provider, its significance lies in its being a set-back on the road to independence. But because the overlapping endpoint is important *for different reasons*, the parties involved may still disagree about the endpoint's degree of significance, or about which interventions to pursue. Thus, the response to divergences, incompatibilities—or even apparent overlap—in research or practice is not monism, it is to consistently ask, "Effective *at what*? *For whom*? *How* effective? And *why that goal*?"

The steps we suggest supplement the compassion and caring already shared by most people involved in SMI research, intervention, and policy. The key motivation for the implementation effort is the promise of progress: Lack of consensus on goals, along with inconsistent or incomplete effectiveness measures, has contributed to the inertia around change and knowledge growth in the US mental health system and elsewhere, and the opposite could move broken systems forward. And if there is one goal that all can agree on, it's that change is needed.

References

1. Abassi, M., Majdzadeh, R., Zali, A., Karimi, A., & Akrami, F. (2018). The evolution of public health ethics frameworks: Systematic review of moral values and norms in public health policy. *Medicine, Health Care and Philosophy, 21*(3), 387–402.

2. American Psychological Association Presidential Task Force on Evidence-Based Practice. (2006). Evidence-based practice in psychology. *American Psychologist, 61*(4), 271–285. https://doi.org/10.1037/0003-066X.61.4.271.
3. Andresen, R., Caputi, P., & Oades, L. G. (2010). Do clinical outcome measures assess consumer-defined recovery? *Psychiatry Research, 177,* 309–317.
4. Arts, B., Simons, C. J., Drukker, M., & van Os, J. (2013). Antipsychotic medications and cognitive functioning in bipolar disorder: Moderating effects of COMT Val108/158 Met genotype. *BMC Psychiatry, 13,* 63. https://doi.org/10.1186/1471-244X-13-63.
5. Bauer, M., El-Khalili, N., Datto, C., Szamosi, J., & Eriksson, H. (2010). A pooled analysis of two randomised, placebo-controlled studies of extended release quetiapine fumarate adjunctive to antidepressant therapy in patients with major depressive disorder. *Journal of Affective Disorders, 127*(1–3), 19–30. https://doi.org/10.1016/j.jad.2010.08.032.
6. Berg, H., & Slaattelid, R. (2017). Facts and values in psychotherapy—A critique of the empirical reduction of psychotherapy within evidence-based practice. *Journal of Evaluation in Clinical Practice, 23,* 1075–1080.
7. Cabassa, L. J., Ezell, J. M., & Lewis-Fernandez, R. (2010). Lifestyle interventions for adults with serious mental illness: A systematic literature review. *Psychiatric Services (Washington, DC), 61*(8), 774–782. https://doi.org/10.1176/ps.2010.61.8.774.
8. Carlier, I. V. E., Meuldijk, D., Van Vliet, I. M., Van Fenema, E., Van der Wee, N. J. A., & Zitman, F. G. (2012). Routine outcome monitoring and feedback on physical or mental health status: Evidence and theory. *Journal of Evaluation in Clinical Practice, 18*(1), 104–110.
9. Carpenter, J. (2002). Mental health recovery paradigm: Implications for social work. *Health and Social Work, 27*(2), 86–94.
10. Cartwright, N., & Stegenga, J. (2011). A theory of evidence for evidence-based policy. In P. Dawid, T. Vasiliki, & M. Vasiliki (Eds.), *Evidence, Inference and Enquiry* (pp. 289–319, Proceedings of the British Academy, Vol. 171): Oxford University Press/British Academy.
11. Chamberlin, J., & Rogers, J. A. (2006). Planning a community-based mental health system, perspective of service recipients. In L. Davidson, C. Harding, & L. Spaniol (Eds.), *Recovery from severe mental illnesses: Research evidence and implications for practice* (Vol. 2, pp. 405–413). Boston: Center for Psychiatric Research.
12. Cipriani, A., Furukawa, T. A., Salanti, G., Chaimani, A., Atkinson, L. Z., Ogawa, Y., et al. (2018). Comparative efficacy and acceptability of 21 antidepressant drugs for the acute treatment of adults with major depressive disorder: A systematic review and network meta-analysis. *The Lancet, 391*(10128), 1357–1366.
13. Davidson, L., & Roe, D. (2007). Recovery from versus recovery in serious mental illness: One strategy for lessening confusion plaguing recovery. *Journal of Mental Health, 16*(4), 459–470.
14. Davidson, L., Harding, C., & Spaniol, L. (2005). Preface. In L. Davidson, C. Harding, & L. Spaniol (Eds.), *Recovery from severe mental illnesses: Research evidence and implications for practice* (Vol. 1, p. xix). Boston: Center for Psychiatric Rehabilitation.
15. Davidson, L., O'Connell, M., Tondora, J., Staeheli, M., & Evans, A. C. (2005). Recovery in serious mental illness: Paradigm shift or shibboleth? In L. Davidson, C. Harding, & L. Spaniol (Eds.), *Recovery from severe mental illnesses: Research evidence and implications for practice* (Vol. 1, pp. 5–26). Boston: Center for Psychiatric Rehabilitation.
16. Douglas, H. E. (2016). Science, policy, values: Exploring the nexus. *Perspectives on Sciene, 24*(5), 475–480.
17. Drake, R. E., Bond, G. R., Goldman, H. H., Hogan, M. F., & Karakus, M. (2016). Individual placement and support services boost employment for people with serious mental illnesses, but funding is lacking. *Health Affairs; Chevy Chase, 35*(6), 1098–1105.
18. Draper, H. (2000). Anorexia nervosa and respecting a refusal of life-prolonging therapy: A limited justification. *Bioethics, 14*(2), 120–133.
19. El-Khalili, N., Joyce, M., Atkinson, S., Buynak, R. J., Datto, C., Lindgren, P., et al. (2010). Extended-release quetiapine fumarate (quetiapine XR) as adjunctive therapy in major depressive disorder (MDD) in patients with an inadequate response to ongoing antidepressant treatment: A multicentre, randomized, double-blind, placebo-controlled study. *International Journal of Neuropsychopharmacology, 13*(7), 917–932. https://doi.org/10.1017/S1461145710000015.
20. Essock, S. M., Goldman, H. H., Van Tosh, L., Anthony, W. A., Appell, C. R., Bond, G. R., et al. (2006). Evidence-based practices: Setting the context and responding to concerns. In L. Davidson, C. Harding,

& L. Spaniol (Eds.), *Recovery from severe mental illnesses: Research evidence and implications for practice* (Vol. 2, pp. 414–434). Boston: Center for Psychiatric Rehabilitation.
21. Frese, F. J., III, Jonathan, S., Kress, K., & Vogel-Scibilia, S. (2006). Integrating evidence-based practices and the recovery model. In L. Davidson, C. Harding, & L. Spaniol (Eds.), *Recovery from severe mental illnesses: Research evidence and implications for practice* (Vol. 2, pp. 375–390). Boston: Center for Psychiatric Research.
22. Garcia, A. M. (2010). What does "work" mean? Reopening the debate about clinical significance. *Clinical Psychology: Science and Practice, 17*(1), 48–51.
23. Green, C. A., Estroff, S. E., Yarborough, B. J. H., Spofford, M., Solloway, M. R., Kitson, R. S., et al. (2014). Directions for future patient-centered and comparative effectiveness research for people with serious mental illness in a learning mental health care system. *Schizophrenia Bulletin, 40*(Suppl 1), S1–S94.
24. Gupta, M. (2013). Psychiatry and evidence-based psychiatry: A distinction with a difference. *Philosophy, Psychiatry, and Psychology, 19*(4), 309–312.
25. Hawthorne, S., & Ihlan, A. (2018). Rethinking civil commitment: The radical resources of the ethics of care. *Public Philosophy Journal*. https://doi.org/10.25335/m5/ppj.1.1-4.
26. Insel, T. R. (2016). *The anatomy of NIMH funding*. https://www.nimh.nih.gov/funding/funding-strategy-for-research-grants/the-anatomy-of-nimh-funding.shtml. Accessed 16 December 2016.
27. Institute of Medicine: Division of Health Promotion and Disease Prevention. (1992). The second fifty years: Promoting health and preventing disability. In R. L. Berg & J. S. Cassells (Eds.), National Academy Press.
28. Ioannidis, J. P. A. (2008). Effectiveness of antidepressants: An evidence myth constructed from a thousand randomized trials? *Philosophy, Ethics, and Humanities in Medicine, 3*(1), 14.
29. Jacobson, N., & Curtis, L. (2006). Recovery as policy in mental health services: Strategies emerging from the states. In L. Davidson, C. Harding, & L. Spaniol (Eds.), *Recovery from severe mental illness: Research evidence and implications for practice* (Vol. 2, pp. 358–374). Boston: Center for Psychiatric Rehabilitation.
30. Jacobson, N. S., Roberts, L. J., Berns, S. B., & McGlinchey, J. B. (1999). Methods for defining and determining the clinical significance of treatment effects: Description, application, and alternatives. *Journal of Consulting and Clinical Psychology, 67*(3), 300–307.
31. Joint Commission on Mental Illness and Health. (1961). *Action for mental health; final report, 1961*. New York: Basic Books.
32. Kessler, R. C., Chiu, W. T., Demler, O., Merikangas, K. R., & Walters, E. E. (2005). Prevalence, severity, and comorbidity of 12-month DSM-IV disorders in the National Comorbidity Survey Replication. *Archives of General Psychiatry, 62*(6), 617–627. https://doi.org/10.1001/archpsyc.62.6.617.
33. Khandaker, G. M., Gandamaneni, P. K., Dibben, C., Cherukuru, S., Cairns, P., & Manaan, K. R. (2013). Evaluating care pathways for community psychiatry in England: A qualitative study. *Journal of Evaluation in Clinical Practice, 19*(2), 289–303.
34. Kincaid, H., Dupre, J., & Wylie, A. (Eds.). (2007). *Value-free science? Ideals and illusions*. Oxford: Oxford University Press.
35. Kravitz, R. L., Duan, N., & Braslow, J. (2004). Evidence-based medicine, heterogeneity of treatment effects, and the trouble with averages. *THe Milbank Quarterly, 82*(4), 661–687.
36. Krimsky, S. (2003). *Science in the private interest: Has the lure of corporate profit corrupted biomedical research?*. Lanham: Rowman & Littlefield.
37. Lambert, M. J., & Ogles, B. M. (2009). Using clinical significance in psychotherapy outcome research: The need for a common procedure and validity data. *Psychotherapy Research, 19*(4–5), 493–501. https://doi.org/10.1080/10503300902849483.
38. Larsen-Barr, M. (2016). *Experiencing antipsychotic medication: From first prescriptions to attempted discontinuation, dissertation*. Auckland: University of Auckland.
39. Leucht, S., Winter-van Rossum, I., Heres, S., Arango, C., & Fleischhacker, W. W. (2015). The optimization of treatment and management of schizophrenia in Europe (OPTiMiSE) trial: Rationale for its methodology and a review of the effectiveness of switching antipsychotics. *Schizophrenia Bullettin, 41*(3), 549–558.
40. Little, M., Lipworth, W., Gordon, J., Markham, P., & Kerridge, I. (2012). Values-based medicine and modest foundationalism. *Journal of Evaluation in Clinical Practice, 18*(5), 1020–1026.
41. Longino, H. E. (1990). *Science as social knowledge: Values and objectivity in scientific inquiry*. Princeton, NJ: Princeton University Press.
42. Longino, H. E. (2002). *The fate of knowledge*. Princeton, NJ: Princeton University Press.

43. Mead, S., & Copeland, M. E. (2005). What recovery means to us: Consumers' perspectives. In L. Davidson, C. Harding, & L. Spaniol (Eds.), *Recovery from severe mental illnesses: Research evidence and implications for practice* (Vol. 1). Boston: Center for Psychiatric Rehabilitation.
44. Morlino, M., Calento, A., Pannone, G., Ventrella, G., & Schiavone, V. (2009). Systematic reviews to support evidence-based psychiatry: What about schizophrenia? *Journal of Evaluation in Clinical Practice, 15*(6), 1025–1028. https://doi.org/10.1111/j.1365-2753.2009.01146.x.
45. Morrison, L. J. (2005). *Talking back to psychiatry: The psychiatric consumer/survivor/ex-patient movement* (New Approaches in Sociology: Studies in Social Inequality, Social Change, and Social Justice, N. Naples, Ed.). New York: Routledge.
46. National Institute of Mental Health. (2017). *Mental illness*. https://www.nimh.nih.gov/health/statistics/mental-illness.shtml. Accessed 17 September 2018.
47. Network, H. V. (2015). *Hearing voices network*. https://www.hearing-voices.org/about-us/real-change-campaign/. Accessed 17 September 2018.
48. Nys, T. R. V., & Nys, M. G. (2006). Psychiatry under pressure: Reflections on psychiatry's drift towards a reductionist biomedical conception of mental illness. *Medicine, Health Care and Philosophy, 9*(1), 107–115.
49. Oaks, D. W. (2011). The moral imperative for dialogue with organizations of survivors of coerced psychiatric human rights violations. In T. W. Kallert, J. E. Mezzich, & J. Monahan (Eds.), *Coercive treatment in psychiatry: Clinical, legal and ethical aspects* (pp. 187–211). Chichester/Hoboken, NJ: Wiley/Blackwell.
50. Padgett, D. K., Stanhope, V., Henwood, B. F., & Stefanic, A. (2011). Substance use outcomes among homeless clients with serious mental illness: Comparing housing first with treatment first programs. *Community Mental Health Journal, 47,* 227–232.
51. Paul, G. (1967). Strategy of outcome research in psychotherapy. *Journal of Consulting Psychology, 31*(2), 109–118.
52. Payne, M. (2016). *Modern social work theory* (4th ed.). New York: Oxford University Press.
53. Pickersgill, M. D. (2014). Debating DSM-5: Diagnosis and the sociology of critique. *Journal of Medical Ethics, 40*(8), 521–525. https://doi.org/10.1136/medethics-2013-101762.
54. Reardon, S. (2017). US mental-health agency's push for basic research has slashed support for clinical trials. *Nature, 546*(7658), 339.
55. Reznek, L. (1991). *The philosophical defence of psychiatry* (Philosophical Issues in Science, ed. W. H. Newton-Smith). London: Routledge.
56. Robotham, D., Wykes, T., Rose, D., Doughty, L., Strange, S., Neale, J., et al. (2016). Service user and carer priorities in a biomedical research centre for mental health. *Journal of Mental Health, 25*(3), 185–188.
57. Rog, D. J. (2005). The evidence on supported housing. In L. Davidson, C. Harding, & L. Spaniol (Eds.), *Recovery from severe mental illnesses: Research evidence and implications for practice* (Vol. 1, pp. 358–376). Boston: Center for Psychiatric Rehabilitation.
58. Shearer-Underhill, C., & Marker, C. (2010). The use of the number needed to treat (NNT) in randomized clinical trials in psychological treatment. *Clinical Psychology: Science and Practice, 17*(1), 41–47.
59. Sisti, D. A., Segal, A. G., & Emanuel, E. J. (2015). Improving long-term psychiatric care: Bring back the asylum (Historical Article). *JAMA, 313*(3), 243–244. https://doi.org/10.1001/jama.2014.16088.
60. Solomon, M. (2015). *Making medical knowledge*. Oxford: Oxford University Press.
61. Stettin, B., Geller, J., Ragosta, K., Cohen, K., & Ghowrwal, J. (2014). *Mental health commitment laws: A survey of the states*. Arlington, VA: Treatment Advocacy Center.
62. Tanenbaum, S. J. (2005). Evidence-based practice as mental health policy: Three controversies and a caveat. *Health Affairs (Millwood), 24*(1), 163–173. https://doi.org/10.1377/hlthaff.24.1.163.
63. Tanenbaum, S. J. (2008). Perspectives on evidence-based practice from consumers in the US public mental health system. *Journal of Evaluation in Clinical Practice, 14,* 699–706.
64. Vanheule, S., & Devisch, I. (2014). Mental suffering and the DSM-5: A critical review. *Journal of Evaluation in Clinical Practice, 20,* 975–980.
65. Wijsbek, H. (2012). 'To thine own self be true': On the loss of integrity as a kind of suffering. *Bioethics, 26*(1), 1–7.
66. Yanos, P. T., Barrow, S. M., & Tsemberis, S. (2004). Community integration in the early phase of housing among homeless persons diagnosed with severe mental illness: Successes and challenges. *Community Mental Health Journal, 40*(2), 133–150.