

Patient-Centered Measurement: Ethics, Epistemology, and Dialogue in Contemporary Medicine

Leah McClimans

N.B. This is the introduction to a book of the same title as the NASPH presentation. The presentation will differ from this introduction, but the latter nevertheless serves as a helpful introduction to the former.

Introduction

Contemporary medicine is Janus-faced. Evidence-based medicine is one face of it, emphasizing evidence, statistics, and method. Patient-centered care is the other, prioritizing patient experiences, judgement, and values. Government agencies, policy makers, major insurers and clinicians have sought ways to bring these faces together. This book is about one such approach, what I call, patient-centered measurement. Patient-centered measures can go by other names: the somewhat cumbersome ‘patient-reported outcome measures’ (PROMs) or the slightly whimsical ‘quality of life measures’. Although I will sometimes use these names when referring to historical moments or well-established measures, in this book I will typically call these instruments ‘patient-centered measures’. This moniker aptly reminds us not only of the tension that exists in the roles these measures are meant to fulfill, but also, I think, better describes the hope they are meant to convey.

Patient-centered measurement is the idea that patient perspectives on, for instance, physical functioning or quality of life, should play an evidentiary role in determining how effective a drug is taken to be, the degree to which a hospital provides good quality care or whether a particular intervention should be funded by an insurer. This idea may sound prosaic, but in fact it’s nothing short of revolutionary. Patient-centered measurement treats patient perspectives *on par* with more traditional metrics such as mortality, morbidity, and safety. It says, patient views matter—not as an afterthought, and not only at the bedside, but in the nuts

and bolts of creating our evidence base, and thus in macro- and meso-level health-care decision-making.¹

What's more, these measures are very popular. They are part of FDA initiatives (FDA 2009; 2023), the UK's development of the NHS (NHS 2023), and Denmark's policy to improve patient care (Egholm et al. 2023). In 2017 the Organization for Economic Co-operation and Development (OECD) published a Ministerial Statement on *The Next Generation of Health Reforms* urging member countries to develop statistical approaches that would allow the assessment and comparability of patient-reported outcomes. They argue use of these outcomes will better equip countries with data on what matters to patients and how well their care is coordinated. Yet despite these policies, initiatives and recommendations, patient-centered measures present a puzzle. And this puzzle has its source in the Janus-faced nature of medicine. How can measurement, which relies on standardization, represent patient perspectives, which, if not idiosyncratic are at least various and changeable?

The stakes for this puzzle are high. If we err on the side of standardization and fail to represent adequately an array of patient perspectives, we risk undermining the ethical values of respect, trust, and autonomy that embody patient-centered care. We also risk epistemic values that speak to our knowledge and understanding of patients' perspectives on, for instance, their functioning or quality of life. If we err on the side of idiosyncratic patient perspectives, we risk the practical, social, and institutional values associated with measuring them, and thus incorporating these perspectives into our evidence base. This book aims to solve this puzzle.

1. Patient-Centered Measures

Before I say more about how I plan to solve this puzzle it might be helpful for some if I say a bit about patient-centered measures, that is, what they look like, how they're used, and their

varieties. Patient-centered measures are presented to people with disabilities, patients, and other ill persons as questionnaires. Some patient-centered measures are ‘generic’, meaning their questions apply to a wide range of illnesses, diseases, and disabilities. For instance, the SF-36 (Ware and Sherbourne 1992) is a well-known generic patient-centered measure that asks people questions like:

Compared to one year ago, how would you rate your health in general now?

- Much better now than one year ago
- Somewhat better now than one year ago
- About the same
- Somewhat worse now than one year ago
- Much worse than one year ago

Other patient-centered measures are condition or disease specific, meaning their questions are restricted to people with a particular illness, disease, or disability. For instance, the Incontinence Impact Questionnaire, Short Form (IIQ-7) (Uebersax et al. 1995), asks women questions like:

Has urine leakage (incontinence) affected your:

Ability to do household chores:

	Not at all	Slightly	Moderately	Greatly
Physical recreation such as walking, swimming, or other exercise:				

Patient-centered measures are largely designed to measure a construct or latent trait.

Although much could be said to problematize the definition, constructs or latent traits are used in this literature to refer to abstract and (currently) unobservable variables such as intelligence, anxiety, subjective physical functioning, or quality of life. Jum Nunnery and Ira Bernstein (1994) have written that constructs reflect a hypothesis that a set of behaviors, say completing household chores independently and taking part in physical recreation, will correlate with one

another or respond similarly to manipulation or intervention. But as I'll discuss in Chapter 1, at least in the context of patient-centered measurement, explicit hypotheses are rare.

Constructs can be multidimensional or unidimensional. Multidimensional constructs have several distinct dimensions related to one another under a single conceptual umbrella. For instance, we might think about anxiety as having cognitive and somatic dimensions. When we think of constructs in relation to measurement, multidimensional constructs are typically measured with multiple scales. Take, for example, the SF-36 (above). The SF-36 measures health status (the construct) on eight different dimensions or scales: physical functioning, role physical, bodily pain, general health, vitality, social functioning, role emotional and mental health. Unidimensional constructs, on the other hand, have a single underlying dimension, for instance, pain is a construct usually taken to be unidimensional. The IIQ-7 (above) is also unidimensional and thus measured on a single scale (Monticone et al. 2020).

When respondents fill out patient-centered measures, they are confronted with a series of questions. In this literature, these questions are often referred to as 'items'. Although I will usually refer to questions as questions, sometimes it will make sense to call them items. And just as questions are sometimes called items, measures are sometimes called 'tests' (even when they aren't what we naturally would think of as a test). I will usually refer to measures as measures, but occasionally it will make more sense to call them tests.

When respondents answer these questions or 'items', their answers are scored to provide measurement or 'test' scores. Some measurement models, for instance, Rasch and item-response theory (IRT), allow for the use of individual scores and population-level scores. Other measurement models like classical test theory (CTT) only allow the use of population-level scores. When a patient-centered measure is multidimensional each scale provides a separate

score. Sometimes—depending on a variety of considerations—those scale scores are presented separately and sometimes they are aggregated and presented as a single score. The patient-centered literature sometimes refers to outcomes instead of scores. The term ‘outcome’ can mean different things within the (vast) measurement literature. Eran Tal (2016), for instance, argues that an outcome is different from the answers or raw scores on a questionnaire. Respondent answers and raw scores, on his account are measurement indications. An outcome is inferred from measurement indications in light of statistical and theoretical assumptions; it is an outcome that refers to a value of the construct, not an indication. But ‘outcome’ also refers to a type of measure used to assess the quality of medical care. In this context an outcome refers to the end point of some process or the effect of some cause. In this context outcome measures are contrasted with process or structural measures (not indications) (Donabedian 2005). Because patient-centered measures are often used as evidence of the quality of medical care, and because in any case they are developed in this milieu, the use of ‘outcome’ often refers to their status as endpoints or effects (for a criticism of this use see McClimans and Browne 2012).

When I first encountered patient-centered measures, they were paper-based questionnaires that had to be mailed to potential respondents and then returned by post. My job was to digitalize these measures by scanning them into a computer. It was a time-consuming and labor-intensive (if boring) process. Now, in many countries, patient-centered measures are completed online. You might find them in your online healthcare portal as a part of measuring your baseline quality of life, and overtime, the impact of an intervention. Or they might be on a device, which a researcher gives you to fill out at various timepoints during a clinical trial. Patient-centered measures are used in a wide variety of contexts from research to regulation to clinical care. But no matter the context of use they share a common purpose: to represent patient

perspectives, to give voice to patients' points of view. But patient points of view are various and changeable, and measurement, at least in this context, tends to expect standardization and invariance. Can patient-centered measures fulfill their purpose? The critical social science literature has long worried about the standardization of patient perspectives and experience (Pols 2006; Mol 2008; Luton 2014). Can these measures overcome those worries? Can this book? I hope so. In the next section, I lay out my plan.

2. Book Outline

Patient-Centered Measurement is the result of over a decade of collaboration with psychometricians, clinicians, epidemiologists, and health services researchers. It's unique, in part, because my experience with these collaborators has provided me with a philosophical analysis of these measuring instruments from the inside out. The puzzle at the heart of it—how to represent adequately patient perspectives in measurement—regularly lurks in the space that surrounds many everyday measurement problems. Health researchers are aware that a tension exists, but they rarely address it head on. Instead, they tend to focus on piece-meal methodological approaches to it, some of which I will discuss in the chapters that follow. As I argue, the result of that approach is patient-centered measures risk forsaking the ethical and epistemic values that make them revolutionary.

Forsaking these values is a genuine problem. It is a problem for patients whose perspectives are potentially co-opted, for the psychometricians, clinicians, epidemiologists, and health services researchers, many of whom get into this research because they believe in the revolutionary nature of these instruments, and for philosophers who have generally ignored or misunderstood the nature of these instruments. What is the solution? *Patient-Centered Measurement* responds with what philosopher's call an 'epistemic' theory drawn from what we

label the hermeneutic tradition, to show how these instruments can be representative and inclusive of patient perspectives. This solution offers an evaluative practice to hone (not a methodology to follow), and a process through which we can improve the epistemic and moral quality of these measures. Words like ‘epistemic’ and ‘hermeneutic’ have precise meanings in philosophy, but I can explain them simply here. Epistemic simply means knowledge, and in this book, it is most usefully contrasted with method. Sometimes methods can give us knowledge, but epistemic investigations, especially in this book, probe the limitations of methods. The hermeneutic tradition—the word derives from the Greek meaning to ‘interpret’—refers to a discipline and a historical movement that treats the interpretation of texts, human action, experience, measurement as a distinct topic of study (George 2021).

In the first two chapters of this book, I discuss how the health science and philosophical literature have failed to address sufficiently the epistemic and ethical values of these instruments. I begin in chapter 1 with the health science literature. This literature is fascinating because on one level researchers are aware of the ethical and epistemic dimensions of their instruments, and, what’s more: it concerns them. For example, researchers and policy makers extol the ethical virtues of patient-centered measures. Their main ethical virtue? To place patients at the center of questions of evaluation. Or as researchers often say to “represent patients’ perspectives”. But they worry about whether their instruments are epistemically up to this task (Fairclough 2017; McKenna 2011; Smith et al. 2019). For instance, in a 2001 article from *The British Medical Journal* titled ‘Are quality of life measures patient-centred?’ (Carr and Higginson 2001) the authors wonder whether standardized metrics can genuinely represent patient perspectives. According to Google Scholar this article has been cited over 700 times. In another article from *The Journal of the American Medical Association*, titled ‘The problem with quality of life in

medicine' (Leplège and Hunt 1997) the authors argue that these measures cannot represent patient perspectives if patients aren't involved in developing the constructs, for instance, physical functioning or quality of life, used by the instruments. This article has been cited over 800 times.

These concerns that I label above as 'epistemic' are developed in the health science literature as concerns over a lack of theory (Hunt 1997; Hobart et al. 2007). Researchers worry that their constructs are theoretically underdeveloped. Moreover, they worry that this lack of theory development jeopardizes the relationship between the measuring instrument and the construct it aims to measure—the relationship that philosophers refer to as coordination, and psychometricians call construct validity. While I agree that patient-centered measurement faces problems with coordination (or validity), I disagree with this diagnosis. Measurement does not require theoretically robust constructs to get started. Rather measuring a construct helps to develop a more theoretically robust construct. For instance, before we had a theory of temperature, early metrologists began measuring temperature by bodily sensations of hot and cold. They noticed that fluids expand their volume when heated, which led to theorizing about why, investigation into heat expansion led to the development of new instruments to measure 'temperature', which eventually led to the development of fixed points, a numerical scale, and quantification. Each innovation increased the scope of metrologist's theorizing of temperature. This example of temperature notwithstanding, as health researchers attempt to solve what they understand as a problem of theory, they tend to reach for off-the-shelf methods and models as the solution. But, as I argue in this chapter, this solution simultaneously complicates and avoids resolving, questions of coordination and validity.

How can we improve the epistemic quality of these instruments, and help them live up to their ethical promise? We need to adopt a lesson learned from studying the history of

measurement in the physical sciences, namely that coordination in measurement has the structure of a hermeneutic circle (Chang 2007; Van Fraassen 2008). The upshot of this lesson is we must accept that we do not have access to value-free, assumption-free, in short, infallible evidence of the constructs we are interested in measuring or their relationship to our measuring instruments. Learning this lesson means embracing a certain amount of uncertainty when we measure and accepting that our methods cannot save us from this fate.

In Chapter 2 I turn from the health science literature to the philosophical literature on well-being. The philosophical literature is almost a mirror image of the health sciences literature—while the health science literature misunderstands the epistemic significance of theory development, the philosophy of well-being literature misunderstands the ethical significance of these measures. Philosophers interested in measurement in the social sciences, tend to assume that the ethical virtue of these instruments is benevolence. That is, they tend to assume that the point of measuring, say, physical functioning or quality of life, is primarily to promote policies that will improve the well-being of populations. Philosophical theories of patient-centered measures are often taken from the well-being literature and are oriented toward this ethical value.

It may be that benevolence is the main ethical virtue in some areas of measurement in the social sciences, for instance, development and economics, but it is not the main ethical virtue in patient-centered measurement. Patient-centered instruments seek to capture patient perspectives and represent patient voices; they are defined by their revolutionary promise to put these perspectives at the center of medical assessment and evaluation. In Chapter 2 I argue that autonomy, not benevolence, best epitomizes the moral impulse of patient-centered measures.

Moreover, I argue specifically that patient-centered measures are *patient-centered* to the degree they:

1. Prioritize patient involvement.
2. Are inclusive of a range of patient perspectives.

In Chapters 3 and 4, I explore these criteria and develop over the course of these chapters an epistemic theory of patient centered measurement. This theory offers a procedural account of patient centered measurement; a procedural account for how to prioritize patient involvement and ensure inclusivity. This account revolves around two processes:

1. Epistemic dialogue.
2. Ongoing coordination.

In Chapter 3 I develop my account of epistemic dialogue. Epistemic dialogue spells out what it means for patient involvement to be prioritized. People with disabilities, patients, and other ill persons should be our primary partners in coming to understand patient-centered constructs and developing their conceptual frameworks. Their testimony provides us with invaluable expertise. It does not mean, however, that health researchers should always simply acquiesce to these understandings. On the contrary, health researchers have a responsibility to understand and operationalize patient-centered constructs, and this responsibility requires asking questions and maybe even disagreeing. The testimony of people with disabilities, patients and other ill persons is invaluable but not infallible.

Here is a preview of my argument. Traditionally, patient-centered measures have been developed by clinicians and health researchers, and then applied to populations of patients. This meant—and sometimes still means—that non-patients develop the questions in a measure that they believe are most important to patients with a particular condition, such as epilepsy. Patients,

then answer these questions, thus providing the information which will form the basis of a claim about the construct. This process of measure development has some fairly obvious problems. It tends, for example, to reinforce identity stereotypes. Here is how it happens. When clinicians and health researchers develop patient-centered measures for people with, for instance, epilepsy they tend to focus on the frequency and strength of seizures because this is what they think is most important to patients' physical functioning or quality of life. People with epilepsy then answer questions about the frequency and strength of their seizures. This information then forms the basis of both 1) a claim about their physical functioning or quality of life and 2) reinforces the claim that frequency and strength of seizures is important to the physical functioning or quality of life of people with epilepsy.

Yet qualitative research involving people with epilepsy suggests that while seizures do affect physical functioning and quality of life, the social stigma of seizures is equally if not more important. Moreover, social support seems to mitigate the anxiety of seizures. The questions clinicians and health researchers ask don't reveal the whole picture of physical functioning or quality of life with epilepsy. Moreover, the part of the picture it does reveal helps to reinforce the stigma that seizures are a barrier to a good life. Relying on clinicians and health researchers to develop a measure's questions in the absence of patients is both epistemically and morally problematic. Epistemically, it limits what we can come to know about, for instance, physical functioning or quality of life in a patient population. Ethically, it is grievously unjust to claim that patients' responses to questions are representative of their perspectives when the questions they are answering may reinforce damaging stereotypes, for instance, people with epilepsy are incapable of employment or violent.

The argument I have outlined so far is broadly similar to themes in the health science literature over the previous decade. Across that period, there was an increasing recognition that representing patient perspectives requires patient representation in measure development, that is, findings from qualitative studies from patients should be used to develop a measure's conceptual model, and the questions used in a measure. Thus, patient-centered instruments should prioritize patient involvement. This requirement has been widely embraced, by the Food and Drug Administration (FDA), the International Society for Pharmacoeconomics and Outcomes Research (ISOPOR), and assorted textbooks (FDA 2009; Rothman et al. 2009; de Vet et al. 2011). Yet this point is still important to make because while the health science literature recognizes the importance of involving people with disabilities, patients and other ill persons in measure development, the philosophical literature, as I will discuss in Chapter 3, has lagged.

In addition to this point about the philosophical literature, it's important to discuss epistemic dialogue in patient-centered measures because epistemic dialogue goes beyond the idea about taking the testimony of people with disabilities, patients, and other ill persons seriously. Epistemic dialogue is not simply a glorified version of qualitative research. Rather it goes beyond qualitative research, and the recommendations from the FDA and ISOPOR to say that just as we must take seriously first-person illness experiences, we must also take seriously the second-hand points of view from health researchers, clinicians and others who have a stake in these instruments. They too have expertise relevant to the development of patient-centered measures. To be sure, it isn't expertise of what it is like to have a particular disability, condition, or illness, but it is still important. Measure development should develop practices and norms that allow health researchers, clinicians and peoples with disabilities, patients, and other ill persons to

engage with one another. If this sounds like pie in the sky philosophy, it isn't. But you'll have to wait for Chapters 3 and 4 to be convinced.

In the development of epistemic dialogue, I conceive of people with disabilities, patients, and other ill persons as partners with health researchers and clinicians in developing patient-centered measures. That is, I conceive of them as having a development role vis-à-vis patient-centered measures. In Chapter 4, I consider people with disabilities, patients, and other ill persons in their role as respondents; as people who answer the questions posed in these measures. In doing so, I address the second criterion for patient-centeredness: inclusion. Let me set the stage by sketching out the problem. Even when people with disabilities, patients, and other ill persons contribute to measure development through epistemic dialogue, it is still the case that the meanings they give questions and answers will change over time (longitudinally) and at a single point in time across individuals (cross-sectionally). In other words, even when we do our very best to create measures with questions that are sensitive to the experience of respondents, these questions won't resonate (or won't resonate in the same way) with all respondents for all time. In a nutshell: patient-centered constructs, and the questions that address them, are context sensitive. If we want our measures to be fit-for-purpose—valid and fair—then we must respond to at least some of these changes in meaning. In other words, we need to ensure our measures are more inclusive of diverse patient experiences. Yet, the questions in these measures are designed to be standardized, not only in their wording but also their meaning.

So, how do we create more inclusive measures? Ultimately, my answer is that patient-centered measures must have ongoing coordination, but to motivate this answer it's useful to begin with some of the literature on what has been terms 'responses shift' or changes in the meaning of a patient's evaluation of the target construct (Sprangers and Schwartz 1999). In the

late 1990's some of the health measurement literature started to recognize that constructs such as quality of life, physical functioning and mobility can have different meanings for different people. The idea that patient-centered measurement has a responsibility to reflect these different meanings is expressed through the literature on individualized quality of life measures (Joyce et al. 1999). These measures forgo attempts to standardize a construct across a population and instead tailor each measure's questions to a particular respondent. The spirit of this concern is advanced through contemporary work on response shift. Unlike individualized quality of life, response shift is often explored in the context of standardized measures. Response shifts are often most visible when a patient responds to the same questionnaire at two or more points in time. Consider, for instance, a person who reports extensive limitations in pursuing leisure activities after one round of chemotherapy, but three months later after five rounds of chemotherapy and worsened health, reports only mild limitations. These answers suggest that quality of life for this person has improved while health has worsened. Examples such as this one have perplexed health researchers. Are response shifts a form of measurement error? Or do they represent important information about respondents' quality of life?

Response shifts are similar in some ways to what philosophers refer to as adaptive preferences. Arguments that explain away changes in preference as merely adaptive tend to assume that the cause of the adaptation, for instance changes in health state due to cancer or disability, is bad (Barnes 2016). But why should we accept this assumption when patients tell a different story—they report that quality of life can improve or remain the same even when health or ability change. Moreover, if response shifts are the result of changes that radically affect respondents in epistemic and personal ways, experiences that L.A. Paul (2014) has called transformative, then it's difficult to know how those of us without the relevant experience can

judge the authenticity of patient responses. In fact, I argue we should not try. We should instead have an inclusive approach, taking patient responses seriously even if (especially if) respondent answers surprise us. This means accepting, in principle, the legitimacy of response shifts, and applying their insights to our measurement models.

In this Chapter 4 I examine two interpretations of what response shift is and how to model it. I examine the ‘appraisal account’ (Schwartz and Rapkin 2004; Rapkin and Schwartz 2004) and the principal of conditional independence (PCI) account (Vanier et al 2021). I argue that, of the two, the appraisal account is better at distinguishing response shift from measurement error. The PCI account says that violations of PCI indicate response shift when observed scores on repeated administrations of a test are not independent of the time of the test (or individual characteristics that change with time) (Oort et al. 2009). What does this mean? It’s helpful to start with PCI. In the context of patient-centered measurement PCI is fulfilled if, given knowledge of, say, physical health, the scores on a test for subjective physical functioning are invariant for, say, gender. Intuitively, this means, considering the physical health of our sample population, knowing their scores on the test tells us nothing about their gender, and knowing their gender tells us nothing about their scores. Violations of PCI that indicate response shift occur when test scores are not invariant with respect to change over time, that is, something about the time between test administrations affects the measurement results.

The appraisal account says that whenever we answer questions, we use a set of cognitive processes to determine the meaning of a question. These cognitive processes are things like establishing a frame of reference (Were you limited in pursuing your leisure activities *with respect to gardening?*) and standards of comparison (Were you *now as opposed to five years ago* limited in pursuing your leisure activities?). Appraisal metrics (Rapkin et al 2017; 2018) help to

measure differences in how respondents understand questions on a patient-centered measure. They provide the *backstory*, if you will, of what is going on when respondents answer questions. Appraisal metrics thus help to interpret respondents' measurement scores and make them more clinically meaningful.

To be sure, not all instances of appraisal are also instances of response shift. Response shifts are essentially unexpected changes in appraisal. But being able to identify response shifts via appraisal is useful because it aids in *coordination*. Coordination refers to the process by which we link measuring processes with constructs. It's through coordination that we can have confidence in a measurement value being a value of a specific construct (instead of something else entirely). Appraisal gives us insight into how respondents understand questions and answers and thus whether their scores meaningfully refer to the construct as it was conceived. Because response shifts will occur, we must anticipate that our measures will lose coordination over time or among different populations. Consequently, measurement coordination is not one and done, it is ongoing. Information from appraisal can feed into coordination in at least two ways. We might use appraisal information to adjust our interpretation of measurement outcomes for some subpopulation. Alternatively, we might alter patient-centered measures themselves to better reflect respondent understandings (or avoid them) or we might redefine the construct of interest—perhaps splitting it into two constructs or reconceptualizing it altogether.

Moving on from Chapters 3 and 4, where I lay out my theory for patient-centered measures, in Chapters 5 and 6 I discuss possible challenges to it. In Chapter 5 I take up the question of measurability of patient-centered constructs, and in Chapter 6 I discuss the pharmaceutical industry's potential to exploit the patient-centered focus I give these instruments.

Are constructs such as physical functioning and quality of life measurable? If this question is posed to psychometricians or philosophers of science, the answers, whether yes or no, will most likely involve a discussion of measurement scales and whether interval level representation is justified. Since I discuss some of these concerns in Chapter 1, in Chapter 5 I set them aside and focus instead on the implications of this question when posed to philosophers of well-being. In the well-being literature, philosophers tend to think of measurability in terms of *heterogeneity* (Hausman 2015; Alexandrova 2017). Heterogeneity is the idea that the goods required to make one individual's life go well are sufficiently different than those required to make other's lives go well. If the goods required for a good life are different for individuals, then measuring a good life is difficult. When you combine heterogeneity with a desire to create measures that are sensitive to individual perspectives, some philosophers of well-being conclude that these constructs aren't measurable or if they are measurable, it's only in a limited way.

In this chapter I compare different accounts of heterogeneity: Dan Hausman's (2015) account in *Valuing Health: Well-Being, Freedom, and Suffering*, Anna Alexandrova's (2017) account in *A Philosophy for the Science of Well-Being* and Bruce Rapkin and Caroline Schwartz and colleagues' account across a number of papers. I argue that all three of these accounts conceptualize heterogeneity as a barrier to measurement, that is, if we are able to measure well-being or quality of life, it's despite heterogeneity. For instance, Hausman and Alexandrova agree that constructs like well-being and quality of life are heterogeneous. For the purposes of using these measures to direct policy, Hausman thinks they are too heterogenous: Alexandrova disagrees. She argues that if we focus on disease and condition specific measures, which she calls contextual well-being, then we can eliminate much of the heterogeneity that makes them problematic. Schwartz et al (2020a), on the other hand, locate heterogeneity not in patient-

centered constructs, but in the perspectives of respondents. If this distinction makes sense (and I argue that it doesn't) it has the benefit of keeping patient-centered constructs homogenous.

Contrary to these approaches, I argue if we can measure patient-centered constructs at all it's because of heterogeneity not despite it. In making this argument I help myself to some lessons from the philosopher Hans-Georg Gadamer's writings on hermeneutics. The upshot of these lessons is that while it's possible to talk abstractly about 'constructs' and 'subjective perspectives', it's only through the application of perspective to constructs, that we have something we can meaningfully study, investigate, probe, or *measure*. In other words, it is only when respondents answer questions about patient-centered constructs that we have empirical content at all. The price of having patient-centered content that we can investigate and measure is that it comes to us as heterogenous material because respondents will often answer questions differently from one another and differently over time. But of course, I don't think this is much of a price because 1) patient-centered measures are an important force for good and 2) we can manage heterogeneity with creative applications of appraisal metrics and ongoing coordination.

In Chapter 6 I turn to a problem with patient-centered approaches and pharmaceutical industries. Recently philosophers have discussed these commercial concerns can co-opt and exploit approaches to incorporate patient perspectives into drug development (e.g. Holman and Geisler 2018; Bueter and Saana 2020). For instance, pharmaceutical companies have sponsored patients to attend FDA patient-focused drug development meetings. The concern is that industry is using the FDA's interest in patient perspectives to capture the discussion at these meetings, and further their own economic interests rather than the interests of patients. Do patient-centered measures play into the hands of industry? Here is the issue. My account of patient-centered measures emphasizes the importance of taking the testimony and responses of people with

disabilities, patients, and other ill persons seriously regardless of “race, religion or creed”. Might this approach to equality open these measures up to exploitation? Moreover, my counterweight to some of these concerns is epistemic dialogue, a form of critical dialogue oriented toward achieving a better—more coherent—understanding through the give and take of questions and answers. But might this approach also be open to capture?

One alternative to taking all patients equally seriously and relying on epistemic dialogue, is to discount the testimony of some patients based on, for instance, their funding or other concerns regarding how they might distort expressions of needs and values or misdirect goals (Warnke 2014). This alternative is what Georgia Warnke (2014) calls “pulling rank”. We might pull rank on some patients because we distrust their testimony for one reason or another. I argue against this alternative for patient-centered measurement, and I argue for further dialogue instead. This isn’t to say that patient-centered measures aren’t open to capture or something similar. In this chapter I offer two examples of how these measures are vulnerable to industry’s economic interests. Yet I also argue that in these cases industry’s ability to manipulate patient-centered measures is largely due to their success in obscuring questions of values (harms and benefits) and economics (profits and losses) while highlighting a narrow version of scientific rigor. This interest in a particular representation of scientific rigor benefits pharmaceutical industries.

Instead of pulling rank, I argue, we ought to change the conversation to include questions of values and economics rather than limit it. In the end, both pulling rank and promoting further dialogue have their risks. But epistemic dialogue avoids the risk of silencing and ignoring patients who should be heard—and who’s to say whose testimony is worthwhile and whose isn’t? History is littered with examples of getting this wrong. Pulling rank, moreover, risks

exacerbating power plays and turning attention away from the things that matter, things like improving health care decision-making and making it more responsive to human needs.

Although I recognize the desire to pull rank, it is not the appropriate response in patient-centered measures: we double down on dialogue.

3. Guide for the Reader

This book tells a story about patient-centered measures with each successive chapter building on the next. At the same time, the chapters have been developed in pairs with each twinset encompassing a theme. Chapters 1 and 2 discuss misconceptions of patient centered measures, first from the context of the health sciences and then from the philosophy of well-being. Chapters 4 and 5 lay out my theory of patient-centered measures explaining how epistemic dialogue prioritizes patients, and ongoing coordination achieves inclusion. In Chapters 5 and 6, I examine challenges to my theory: measurability and industry exploitation. Readers particularly interested in one or more of these themes can pick up these chapters out of order. A final note. About half-way through writing this book I bought a horse and then...a pony. Examples and stories of my adventures pepper these chapters beginning with Chapter 4, so if you love horses you may want to start here.

References

Alexandrova, Anna. 2017. *A Philosophy for the Science of Well-Being*. Oxford: Oxford University Press.

Barnes, Elizabeth. 2016. *The Minority Body*. Oxford: Oxford University Press.

Bueter, Anke and Saana Jukola. 2020. 'Sex, Drugs, and How to Deal with Criticism: The Case of Flibanserin'. In *Uncertainty in Pharmacology, Epistemology, Methods and Decisions*, edited by Adam LaCaze and Barbara Osimani, pp. 451-69. Switzerland: Springer Nature.

- Carr, Allison J. and Irene J. Higginson. 2001. 'Are quality of life measures patient centred?' *British Medical Journal* 322: pp. 1357-60.
- de Vet, Henrica C.W., Caroline B. Terwee, Libwine B. Mokkink, and Dirk L. Knol. 2011. *Measurement in Medicine: A Practical Guide*. Cambridge: Cambridge University Press.
- Donabedian, Avedis. 2005. 'Evaluating the Quality of Medical Care'. *The Milbank Quarterly* 83 (4): 691-729.
- Egholm, Cecilie Lindström, Sanne Jensen, Annette Wandel, and Mogens Hørder. 2023. 'The implementation of the 2017 national policy on patient-reported outcomes in Denmark: An overview of developments after six years'. *Health Policy* 130: 104755.
- Fairclough, Diane L. 2017. 'Challenges of interpreting patient reported outcomes from clinical trials'. *Annals of Translational Medicine* 5 (2): 408.
- FDA. 2009. 'Guidance for Industry on Patient-Reported Outcomes Measures: Use in medicinal product development to support labeling claims'. *Federal Register* 74: pp. 1-43.
- FDA. 2023. 'FDA Patient Focused Drug Development Guidance Series for Enhancing the Incorporation of the Patient's Voice in Medical Product Development and Regulatory Decision Making'. Last modified 6 April 2023. <https://www.fda.gov/drugs/development-approval-process-drugs/fda-patient-focused-drug-development-guidance-series-enhancing-incorporation-patients-voice-medical>.
- George, Theodore. 2021. 'Hermeneutics', *The Stanford Encyclopedia of Philosophy*. Edited by Edward N. Zalta. Accessed 22 June 2023. URL = <https://plato.stanford.edu/archives/win2021/entries/hermeneutics/>.

- Hasok, Chang. 2004. *Inventing Temperature: Measurement and Scientific Progress*. Oxford: Oxford University Press.
- Hausman, Daniel M. 2015. *Valuing Health Well-Being, Freedom, and Suffering*. Oxford: Oxford University Press
- Hobart, Jeremy C., Stefan J. Cano, John P. Zajicek, and Alan J. Thompson. 2007. 'Rating Scales as Outcome Measures for Clinical Trials in Neurology: Problems, Solutions, and Recommendations'. *Lancet Neurology* 6: pp. 1094-105.
- Holman, Bennet and Sally Geislar. 2018. 'Sex, Drugs and Corporate Ventriloquism: How to Evaluate Science Policies Intended to Manage Industry-Funded Bias'. *Philosophy of Science* 85, no. 5: pp. 869-81.
- Hunt, Sonja M. 1997. 'The problem of quality of life'. *Quality of Life Research* 6: pp. 205-212.
- Uebersax, John S., Jean F. Wyman, Sally A. Shucker, and Donna K. McClish, and Andrew Fantl. 1995. 'Short forms to assess life quality and symptom distress for urinary incontinence in women: the Incontinence Impact Questionnaire and the Urogenital Distress Inventory. Continence Program for Women Research Group'. *Neurourology and Urodynamics* 14 (2): pp. 131-9.
- Joyce, Charles Richard Boddington, Hannah M. McGee, and Ciaran A. O'Boyle. 1999. *Quality of Life in Individual Quality of Life: Approaches to Conceptualization and Assessment*. Harwood Academic Publishers.
- Leplège, Alain and Sonia Hunt. 1997. 'The Problem of Quality of Life in Medicine'. *Journal of the American Medical Association* 278 (1): pp. 47-50.
- Lupton, Deborah. 2014. "The Commodification of Patient Opinion: The Digital Patient Experience Economy in the Age of Big Data." *Sociology of Health & Illness* 36 (6): 856–69.

- McClimans, Leah and John Browne. 2012. Quality of Life is a Process Not an Outcome. *Theoretical Medicine and Bioethics*, 33: 279-92.
- McKenna, Stephen. 2011. "Measurement patient-reported outcomes: moving beyond misplaced sense to hard science". *BMC Medicine* 9: 86.
- Mol, Annemarie. 2008. *The Logic of Care: Health and the Problem of Patient Choice*. London and New York: Routledge.
- Monticone, Marco, Giorgio Ferriero, Andrea Giordano, Calogero Foti, and Franco Franchignoni. 2020. 'Rasch analysis of the Incontinence Impact Questionnaire short version (IIQ-7) in women with urinary incontinence'. *International Journal of Rehabilitation Research* 43 (3): pp. 261-56.
- NHS. 'Patient Reported Outcome Measures (PROMS)'. Accessed 17 June 2023. <https://www.england.nhs.uk/statistics/statistical-work-areas/proms/>.
- Nunnery, Jum C and Ira H. Bernstein. 1994. *Psychometric Theory*, 3rd ed. New York: McGraw-Hill.
- The Organization for Economic Cooperation and Development. 2017. Ministerial Statement *The Next Generation of Health Reforms*. Accessed 20 June 2023. <https://www.oecd.org/newsroom/oecd-health-ministerial-statement-the-next-generation-of-health-reforms.htm>
- Paul, Laurie A. 2014. *Transformative Experience*. Oxford: Oxford University Press
- Pols, Jeannette. 2006. "Accounting and Washing: Good Care in Long-Term Psychiatry." *Science, Technology, & Human Values* 31 (4): 409–30.
- Rapkin, Bruce and Carolyn Schwartz. 2004. 'Toward a theoretical model of quality-of-life appraisal: Implications of findings from studies of response shift'. *Health and Quality of Life Outcomes* 2: pp. 14.

Rapkin, Bruce, Iliana Garcia, Wesley Michael, Jie Zhang and Carolyn Schwartz. 2017.

‘Distinguishing appraisal and personality influences on quality of life in chronic illness: introducing the quality-of-life Appraisal Profile version 2’. *Quality of Life Research* 26 (9): pp. 2815-29.

Rapkin, Bruce, Iliana Garcia, Wesley Michael, Jie Zhang, and Carolyn Schwartz. 2018.

‘Development of a practical outcome measure to account for individual differences in quality-of-life appraisal: the Brief Appraisal Inventory’. *Quality of Life Research* 27 (3): pp. 823-33.

Rotherman, Margaret, Laurie Burke, Pennifer Erickson, Nancy Kline Leidy, Donald L. Patrick,

and Charles D. Petrie. 2009. ‘Use of existing Patient-Reported Outcome (PRO) instruments and their modification: The ISPOR good research practices for evaluating and documenting content validity for the use of existing instruments and their modifications PRO Task Force Report’. *Value in Health* 12 (8): pp. 1075-83.

Schwartz, Carolyn and Bruce Rapkin. 2004. ‘Reconsidering the psychometrics of quality of life

assessment in light of response shift and appraisal’. *Health and Quality of Life Outcomes* 2, no. 1: pp. 16.

Schwartz, Carolyn, Roland Stark, and Bruce Rapkin. 2020a. ‘Capturing patient experience: does

quality of life appraisal entail a new class of measurement?’ *Journal of Patient-Reported Outcomes* 4 (1): pp. 85.

Smith, Sarah, Stefan Cano, and John Browne. 2019. ‘Patient reported outcome measurement:

drawbacks of existing methods’. *British Medical Journal* 364: pp. 1844.

Sprangers, Mirjam and Carolyn Schwartz. 1999. 'Integrating response shift into health-related quality-of-life research: A theoretical model'. *Social Science & Medicine* 48, no. 11: pp. 1507-15.

Tal, Eran. 2016. 'Making Time: A Study in the Epistemology of Measurement'. *British Journal for the Philosophy of Science* 67 (1): pp. 297-335.

Van Fraassen, Bas. 2008. *Scientific Representations: Paradoxes of Perspective*. Oxford: Oxford University Press.

Vanier, Antoine, Frans Oort, Leah McClimans, Nikki Ow, Bernice Gulek, Jan Böhnke, Mirjam Sprangers, Véronique Sébille and Nancy Mayo. 2021. 'Response shift in patient-reported outcomes: definition, theory, and a revised model' *Quality of Life Research* 30 (12): pp. 1-14.

¹ Patient-centered measures are also used increasingly in the clinical arena where they can be implemented to improve the individualization of patient care, and patient-clinician communication.