Note: These are the first two chapters of my book manuscript on well-being. My talk, "The nature of well-being", will cover this material plus some more material from my book. But these two chapters should give you a good sense of what I'll be arguing for.

Chapter 1. A Short Introduction to the Causal Network View of Well-Being

The causal network view of well-being, while novel and perhaps unusual, is at bottom simple and intuitively appealing. My goal in this chapter is to describe this view as I might to a friend or sibling: simply, succinctly and with no theoretical fuss. I will not try to satisfy the nattering critic that sits on my shoulder, or yours. We'll have the rest of the book to deal with them.

A good way to begin to think about well-being is with an exercise. How would you explain that a person has a high degree of well-being without actually using the word 'well-being' or its synonyms? If you aren't already corrupted by a philosophical theory, you might offer a thumbnail sketch something like this: "Susan is in a happy and fulfilling committed relationship, she has close and caring friends, she keeps fit by playing racquetball, a sport she enjoys, and her professional life is both successful and satisfying." Most people's theoretically innocent description of someone with a high degree of well-being will include both "objective" facts about the person (e.g., facts about her relationships, her activities, her professional life) as well as "subjective" facts about the person (e.g., facts about her commitments to and her feelings about kith and kin, facts about her finding certain activities enjoyable and satisfying). The objective and subjective facts we appeal to when explaining Susan's – or anyone's – degree of well-being include:

1. positive feelings, moods, emotions (e.g., joy, contentment),
2. positive attitudes (e.g., optimism, hope, openness to new experiences),
3. positive behaviors and traits (e.g., friendliness, curiosity, perseverance), and
4. successful interactions with the world (e.g., strong relationships, professional accomplishment, fulfilling hobbies or projects).

So far, so good. But how does this ramshackle set of facts fit into a coherent whole? How are we supposed to unite these various objective and subjective facts into a coherent theory of well-being? The answer I propose is simple: We don't have to. That's because the world has already joined them together with causal bonds. Think of the above elements of someone's well-being as nodes in a complex causal network or web. Every node in that network is causally connected to some of the other nodes – it fosters some, and is fostered by others. A person who has well-being is in a "positive rut" – she is enmeshed in a positive causal web involving positive feelings, attitudes, behaviors, traits and successful interactions with the world. Well-being is just this state of being in this positive causal network, nothing more and nothing less. To construct a theory of well-being, the only theorizing we need to do is to account for this type of state.

There is a way to explain the gist of this view that appeals to commonsense: Well-being is being in a "success breeds success" cycle. Of course, it would be just as accurate and incomplete, though considerably less pithy, to say that well-being is being in a "positive feeling breeds positive feeling" cycle, or a "positive attitude breeds positive attitude" cycle, or a

"positive behavioral disposition breeds positive behavioral disposition" cycle. To have a high degree of well-being is to instantiate a causal web involving all of these sorts of states.

We can put some meat on this somewhat abstract discussion by turning again to the features of Susan's life we appeal to in explaining her well-being – her committed relationship, her friendships, her exercise regimen, her professional success, her confidence and sense of mastery, her joie de vivre, her friendliness, her moxie and adventurousness, her curiosity, her hope and optimism. This is not an accidental conglomeration of happy facts. Susan's well-being has a causal structure – the components of her well-being are causally bound together. Pick out any "happy fact" that is a component of Susan's well-being, say, her professional success. This success is fostered by many other factors we appeal to in describing Susan's well-being – her curiosity, moxie, optimism, and confidence, her exercise regimen, her social support. But Susan's professional success is also a cause of some of those "happy facts." Her professional success bolsters her income, her optimism, her confidence and the strength of her relationships. So Susan's professional success is a node in a causal network of factors that constitute a significant portion of her well-being. Susan's professional success is a node in a positive causal web.

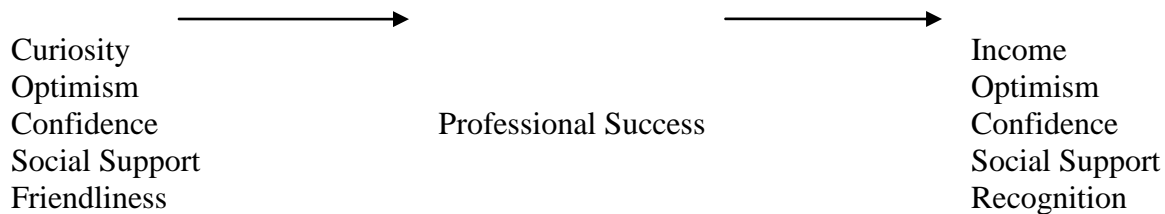| Curiosity | | Income |
Optimism		Optimism
Confidence	Professional Success	Confidence
Social Support		Social Support
Friendliness		Recognition

Figure 1: Professional Success as a Node in a Positive Causal Network

What is true about Susan's professional success is also true of many of the components of her well-being. Each of these is a node in a causal network of some combination of positive feelings, positive attitudes, positive traits or successful interactions with the world.

An important feature of figure 1 is that some states (e.g., Susan's optimism, confidence, and social support) are both cause and effect of Susan's professional success. What this suggests is that Susan's well-being consists of some positive cyclical processes. Susan's professional success leads her to acquire, maintain or strengthen other positive features of her person; and in turn these positive features help foster her professional success; and so on (figure 2). Positive cycles are plausibly associated with many other components of people's well-being. For example, Susan's optimism helps her overcome challenges and makes her more successful socially and professionally, and having success tends to bolster Susan's optimism (Seligman 1990). Susan's friendships and committed relationship provides her with various kinds of material and psychological support, which help to make Susan more trusting, more extraverted, more generous, and these traits in turn make Susan a better friend and partner, which tend to strengthen her friendships and relationship. Susan's exercise regimen gives her more strength, energy and positive emotions, which help her, in various ways, to continue her exercise regimen. And so on.

Optimism
Confidence
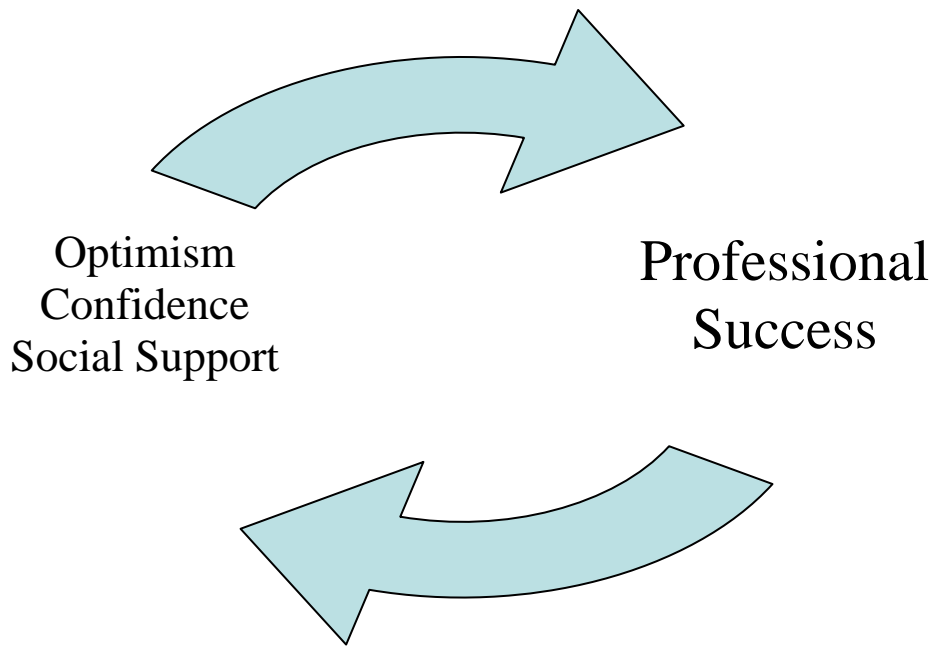Social Support

Professional
Success

Figure 2. The Positive Cycle of Professional Success

A useful way to understand positive cycles is to contrast them with negative or vicious cycles. Think about a time in your life when you were "blue" or "down." Even if your episode did not meet the DSM (The Diagnostic and Statistical Manual of Mental Disorders) criteria for an affective disorder, any memorable blue period will include some of the characteristic features of dysthymia – depressed mood, diminished interest or pleasure in activities, loss of energy, feelings of worthlessness, excessive or inappropriate guilt, morbid thoughts, indecisiveness, diminished ability to think or concentrate, insomnia or hypersomnia, psychomotor agitation or retardation, and some "impairment in social occupational, or other important areas of functioning" (200x, yy). What is cruel about dysthymia, and what you might have noticed during your blue period, is that these criteria are links in self-maintaining causal cycles. Negative feelings, attitudes, behaviors and dysfunctions causally build upon and reinforce one another. A deep melancholy can lead to life problems (at school, at work, in friendships and relationships), and these life problems can in turn produce greater feelings of melancholy and despair, which can in turn produce more life troubles, etc. Anyone who has suffered through a deep blue period, much less a serious depression, will not need to be convinced of the grim reality of these sorts of vicious cycles. Positive cycles can be plausibly understood as the mirror images of these vicious cycles.

Let's review. The causal network (CN) view of well-being maintains that positive networks and cycles are the basic building blocks of well-being. A positive network is a causal network involving positive feelings, emotions, attitudes, behaviors, traits, and successful interactions with the world (figure 1). A positive cycle is a special kind of positive network - one that tends to cycle back over the same states or conditions (figure 2). A person's well-being is nothing more than her manifesting a robust set of positive causal networks and positive cycles. Susan having a high degree of well-being is simply Susan being caught up in a rich set of

powerful positive networks and cycles. Before going into more detail about this view, I want to motivate it by highlighting some features that, I think, make it attractive.

1. Well-being is a real condition. The CN view yields the plausible result that well-being is a real state or condition people can be in. It is as real as being depressed or having the flu. What's more, well-being seems to involve both "subjective" facts (e.g., positive affect) and "objective" facts (e.g., close friends, professional success) about a person. A challenge for any theory of well-being is how to stitch these disparate facts together into a coherent whole. CN connects the objective and subjective components of well-being in a simple and natural way, in the way the world connects them: with causal bonds.

2. Well-being is reasonably causally stable. The CN view is in a position to smoothly explain why well-being is causally stable, to the extent it is causally stable: It is a condition that tends to be self-reinforcing or self-perpetuating. Well-being will be a self-maintaining "rut" whenever the causal networks that make it up consist of many positive cycles. In Susan's case, for example, if there are positive cycles that maintain and reinforce her professional success, her positive affect, her confidence, her optimism, her strong relationships, etc., then her well-being will be a robust, stable condition. It goes without saying that well-being and its component processes are neither permanent nor inevitable. There are many factors, both environmental and internal to the person, that can frustrate or extinguish these positive cycles. And if they do enough damage to those positive causal processes, they can undermine well-being. But the reason well-being tends to be a self-maintaining sort of condition is because the causal network that is well-being tends to consist of positive cycles.

3. Well-being is multiply realizable. The CN view accounts for the fact that well-being can be instantiated in many different ways. The well-being of the professional athlete and the well-being of the spritely octogenarian can be made up of quite different characteristics. CN can provide a unified account of this multiply realizable state – in every case, well-being involves a dynamic state involving self-maintaining causal links between positive feelings, emotions, attitudes, traits and interactions with the world. But what precisely those links are can vary quite a bit from person to person.

4. Explaining the complex relations between well-being and its constituents. Let's focus in more detail on the relationship between professional success and well-being. In the example above, professional success is a constituent of Susan's well-being. But we need not assume that Susan's professional success is causally related only to good things. Anyone's professional success, including Susan's, will typically be causally connected to some stress, conflict and unhappiness. Some moderate degree of struggle and woe is perfectly consistent with, and in some cases perhaps even essential to, having a high degree of well-being. The CN view can account for this smoothly: A full professional life will inevitably entail overcoming challenging obstacles and difficulties, which will include some measure of failure and negative feelings. But as long as those negative facts don't destroy or destabilize the positive causal network partially represented in figure 1 (which includes as nodes Susan's professional success, confidence, deep satisfaction, optimism, social support and recognition), Susan's professional success will remain a component of Susan's well-being. What's more, it is not outrageous to suggest that these negative facts (the failures and negative feelings) might ultimately strengthen that positive causal web. The fact that

Susan has overcome such difficulties can contribute to her optimism and confidence; the fact that others see that she successfully resolves difficult issues at work can give her greater recognition and social support. So it is possible that those negative facts are essential to Susan's level of well-being: Susan has a higher degree of well-being than she would have had she not persevered through those challenges with their attendant failures and negative feelings, moods and emotions.

The CN view can also smoothly account for the fact that even though professional success is part of Susan's well-being, it is neither necessary nor sufficient for well-being in general. For example, Mark might not have much professional ambition or success and still have a high degree of well-being. He might find himself enmeshed in many positive networks and cycles involving his personal projects, his relationships with kith and kin, and still not have much professional success. In this way, professional success is not necessary for well-being. Nor is it sufficient. Consider George, who despite his professional success leads a life with no significant amount of well-being. The CN view allows for this possibility: George's professional success is not enmeshed in a positive causal web. Instead, his professional success is brought about by unhealthy psychological insecurities and obsessions and produces stress, unhappiness and self-loathing. And unlike Mark, George does not instantiate any positive networks and cycles involving personal projects, friends or family. And so in this way, the CN view is consistent with the truism that professional success is not sufficient for well-being. But there's a more subtle possibility readily explained by the CN view of well-being: Someone can have well-being and professional success without that success contributing to her well-being. Suppose Sandra, like Mark, finds herself enmeshed in many positive ruts involving her hobbies and her relationships with kith and kin. These positive ruts are robust and strong, and so Sandra, like Mark, has a high degree of well-being. But unlike Mark, Sandra is highly competent at her job and has considerable professional success. But Sandra does her work from a sense of duty and while she does not find it unpleasant, she gets no particular joy or satisfaction from it either. Sandra's professional success is causally isolated from the networks that constitute Sandra's well-being - it is a causal dangler. It lacks significant causal connections to any rich set of Sandra's positive feelings, emotions, attitudes, or traits. And so it does not contribute to Sandra's well-being.

5. CN explains what's right and what's wrong with many traditional theories of well-being. This issue will be the topic of chapter 4. But it is worth mentioning it briefly in this introduction. While philosophers have defended theories of well-being, it is fair to say that no theory is widely accepted and every theory is generally thought to suffer from some serious deficits. I contend that the CN view explains much of both what is right and what is wrong about these theories. Consider, for example, desire satisfaction views of well-being (which hold, roughly, that well-being is getting what you want). Take any desire George might have – no matter how strong or well-informed that desire might be. It is always an open empirical question how the satisfaction of that desire will affect the positive causal processes that constitute (or that might constitute) George's well-being. Suppose George satisfies a desire that results in his spending time on a project that brings no joy, no satisfaction, no recognition. Or worse, suppose that the satisfaction of the desire results in his spending time on a project that brings grief and derision. There are many things one might say about that desire - it might have been exceptionally strong or fully informed, or it might have satisfied an important moral imperative. But if the satisfaction of that desire undermined the smooth operation of a significant portion of the positive causal networks George instantiates (which on the CN view make up George's well-being), then it would have

undermined his well-being. The connection between desire-satisfaction and improved well-being is a contingent one. And so any attempt to argue that desire satisfaction is constitutive of well-being will fail. Nonetheless, there is power in the desire theory, and I submit that the CN view explains its power. Our desires, when properly informed, will tend to be for those things that establish or maintain our positive causal networks (or at least support the conditions needed for those networks to be established and maintained). And so well-being does typically involve the satisfaction of our properly informed desires.

At this point, the nattering critic on my shoulder – and I suspect the nattering critic on your shoulder – will not be silenced. There are many questions and objections still to answer about this view of well-being. For example:

- What evidence is there for believing the CN view of well-being?
- What is it for a feeling, emotion, attitude, or trait to be <u>positive</u>? And what is it for an interaction with the world to be <u>successful</u>?
- How can the CN view account for the graded nature of well-being?
- Can the CN view account for the normativity of well-being?

These difficult questions demand considerable attention, and so they fall outside the scope of this short, intuitive introduction. I will address them in upcoming chapters.

Chapter 2. How to Investigate the Nature of Well-Being

What we want from a theory of well-being is straightforward enough: We want a theory that accurately and perspicuously describes the nature of well-being. But how are we supposed to arrive at such a theory? What evidence should we consider? And how should we proceed from that evidence? Getting clear about these questions is important for two reasons. First, our methods demarcate the sorts of considerations that we take to be legitimately raised for or against a proposed theory of well-being. And second, the methods we adopt inevitably bind us to substantive assumptions about the nature of well-being and how to investigate it. We adopt one method rather than another because (presumably) we think that there are better and worse ways to find out about well-being and <u>this</u> way of finding out about well-being is a particularly effective one. Of course, whether some particular method really is a good one for learning about well-being depends on what well-being is like. A form of Meno's paradox looms: To discover the nature of well-being, we need a plan or a method to find out about it; but such a method must presuppose something about the nature of well-being.

Today, there is no widely-accepted, powerful theory of well-being akin to, say, Newton's physics in the early 19th century. In absence of such a theory, it might seem unwise for philosophers to put all our methodological eggs in one basket. The discipline of philosophy maximizes its chances of developing a powerful theory if philosophers adopt a range of different methods for studying well-being (Kitcher 1990). It is unfortunate, therefore, that philosophers's efforts to investigate the nature of well-being are dominated by a single basic method. After all, no matter how "obvious" any approach might seem to us, there is always the risk that it will lead to a dead end. (On some tellings, this is what happened to logical empiricism – it was felled, not by some powerful, dramatic argument to the effect that its central pillars were false, but by its inability, over decades, to deliver on its promises.) My goal in this chapter is to raise some worries about the traditional philosophical approach to the study of well-being and to propose an alternative. The alternative method I propose is (as far as I know) new to the study of well-being. But the method makes philosophical assumptions that are widely held and prima facie plausible. It is familiar in many other areas of philosophy. Nonetheless, there are no sound a priori arguments showing that some proposed approach must bear fruit. Every approach must ultimately judged by whether it delivers a promising theory.

1. The Traditional Philosophical Approach, Part I: Descriptive Adequacy

The traditional philosophical approach to the investigation of well-being consists of two elements: the Descriptive Adequacy condition and the Philosophy First assumption. Let's begin with the Descriptive Adequacy condition, which every philosopher I am aware of who has written about well-being adopts (either explicitly or implicitly by how they go about investigating well-being). It says that by far the most important source of evidence that any theory of well-being must account for is our ordinary, commonsense judgments about well-being. James Griffin states that "the notion we are after is the ordinary notion of 'well-being'" (1986, 10). In L.W. Sumner's discussion of what we want from a theory of welfare, which he takes to be "more or less the same" as well-being (1996, 1), he concisely articulates the first part of the traditional approach:

> [T]he best theory about the nature of [well-being] is the one which is most faithful to our ordinary concept and our ordinary experience. That experience is given by what we think or feel or know about well-being, both our own and that of others. The data which a candidate theory must fit, therefore, consist of the prodigious variety of our preanalytic convictions (1996, 10-11).

The best theory of well-being is the one that "makes the best sense of [our preanalytic] convictions" (1996, 11). The degree of fit between a theory of well-being and our pretheoretic convictions is "a function of the extent to which the truth conditions [the theory] offers can support and systematize our intuitive assessments." While Sumner does not argue that "descriptive adequacy" is the sole requirement that the correct theory of well-being must satisfy, it is clearly the most important – it is "the basic test" (1996, 10).

Valerie Tiberius argues that "[t]o give an analysis of the nature of well-being is to articulate the criterion (or criteria) that any thing must meet in order to count as a source or cause of well-being" (2004, 295-6). Like Sumner, she argues that such theories are to be evaluated according to the Descriptive Adequacy condition.

> Formal analyses [which provide an account of the nature of well-being] are to be evaluated on the basis of how well they accommodate our uses of the concept in question and how well they fit with our ordinary experience. In other words, formal accounts of well-being are evaluated primarily in terms of their descriptive adequacy. The most descriptively adequate account of well-being is the one that is most faithful to our pre-philosophical convictions about well-being (299).

So a theory that tells us about the nature of well-being must be "faithful" to our commonsense judgments about well-being.

The Descriptive Adequacy condition is ubiquitous in the philosophical literature on well-being. It is implicit in every one of the currently "standard" counterexamples to the dominant theories of well-being. For example, hedonism holds that a person's well-being is a function of the balance of her positively valenced experience (e.g., enjoyment, pleasure, happiness) over her negatively valenced experience (e.g., suffering, pain). A well-known counterexample to hedonism asks us to consider two people who have exactly the same experiences, but one is genuinely engaged with the world and the other is prone in a laboratory hooked up to an Experience Machine (Nozick 1974). Hedonism implies that they have the same degree of well-being. But because commonsense tells us that the genuinely engaged subject has a greater degree of well-being, this example is taken to be a powerful reason to reject hedonism. The Descriptive Adequacy condition is implicit in this rejection: Because hedonism is inconsistent with these convictions, with "what we think or feel or know about well-being" (Sumner 1996, 10), it cannot be the correct view about the nature of well-being.

2. The Traditional Philosophical Approach, Part II: Philosophy First

The Descriptive Adequacy condition holds that our convictions about well-being serve as the primary arbiters, the primary evidence, for our theories of well-being. "We manifest these

convictions whenever we judge that our lives are going well or badly, that pursuing some objective will be profitable or advantageous for us, that a change in our circumstances has left us better or worse off, that some policy would enhance or erode our quality of life, that some measure is necessary in order to protect the interest of our family or community, that a practice which is beneficial for us may be harmful to others, that we are enjoying a higher standard of living than our forebears, and so on" (Sumner 1996, 11). And yet, Sumner rightly insists that philosophers have "no special expertise" when it comes to advising "us what is good or bad for us, or [advising] us on how to attain the former and avoid the latter" (1996, 7). There is a prima facie tension here. On the one hand, philosophers have "no special expertise" about "what is good or bad for us." And on the other hand, philosophers' convictions implicit in judgments like "our lives are going well or badly" are so belief-worthy that they constitute the primary evidence for theories of well-being. How can this be?

The solution to this tension involves distinguishing between judgments about the <u>nature</u> of well-being and judgments about the <u>causes, effects and correlates</u> of well-being. Consider the following four judgments.

5. Ceteris paribus, well-being is undermined when someone hooks up to the Experience Machine.
6. Ceteris paribus, well-being is undermined when someone has false friends who seem genuine.
7. Ceteris paribus, well-being is undermined when someone's desire to count the blades of grass on the college lawn is satisfied.
8. Ceteris paribus, well-being is undermined when someone has a longer commute (e.g., has a larger house in the suburbs rather than a smaller house in the city, nearer one's place of work) (citation).

Philosophers wedded to the traditional approach have rejected theories of well-being that do not accord with the first three judgments. So in practice, the traditional approach is committed to the proposition that the first three judgments are about the nature of well-being and constitute evidence that a theory of well-being must make sense of. (That's not to say these judgments are indefeasible, see Crisp 2006.) But the fourth judgment is different. Even if it is true – and even if we have overwhelming evidence that it is true – it is not a judgment about the <u>nature</u> of well-being. Rather, it is an empirical judgment about what tends to foster or undermine well-being. And so according to the traditional approach, it is a judgment that does not form the primary evidential base for a theory of well-being. Therefore, proponents of the traditional approach can admit that the fourth judgment is one about which philosophers have "no special expertise." (It is natural to suppose that the traditional approach assumes that our judgments about the nature of well-being are a priori and that our judgments about the causes, effects and correlates of well-being are a posteriori. And that is why philosophers can have expertise about the former but not the latter. But some proponents of the traditional approach might reasonably deny that they are committed to this distinction. So for the purposes of capturing as wide a range of philosophers as possible, I will not assume that this epistemological thesis is part of the traditional approach.)

This distinction between judgments about the nature of well-being and judgments about the causes and correlates of well-being leads ineluctably to a Philosophy First approach to

investigating the nature of well-being. The intuitive idea is that the philosophical project of accounting for the nature of well-being is foundational and logically prior to the empirical project of identifying the causes, effects and correlates of well-being. More specifically, the Philosophy First approach involves two commitments:

· Insulation: Philosophy is insulated from psychology. The primary evidence for or against a theory about the nature of well-being comes from our judgments about the nature of well-being. Empirical findings about the causes, effects or correlates of well-being are not relevant evidence for or against a theory about the nature of well-being.
· Vulnerability: Psychology is vulnerable to philosophy. Any empirical claim about well-being must make substantive assumptions about the nature of well-being that are ultimately validated or invalidated by the correct philosophical theory about the nature of well-being.

The philosophical project is foundational. Psychology cannot safely ignore philosophy; but philosophy can safely ignore psychology.

We can appreciate the power of the Philosophy First approach with an example. Some psychologists have argued that meditation tends to foster well-being (citation). Since these psychologists are not measuring well-being directly, their evidence for this hypothesis is indirect. So let's suppose that it is true that meditation tends to foster X, where X is an empirical property and these psychologists have taken X to accurately measure well-being (or a proxy for well-being). We can distinguish two claims:

1. Meditation tends to foster X.
2. Meditation tends to foster well-being.

(1) is the sort of claim about which philosophers have no special expertise. Let's assume that it's true. What about (2)? Whether (2) follows from (1) depends on whether psychologists are right to suppose that X is a good measure or proxy for well-being. But this will largely depend on what well-being is; and (according to the traditional approach) this is something we can discover only via philosophy, via reflection on our judgments about the nature of well-being. An accurate account of the nature of well-being will be delivered by the correct <u>philosophical</u> theory of well-being. The argument for the Philosophy First approach comes in the form of a dilemma: The assumption that X is a good measure of well-being is either consistent with the correct philosophical theory of well-being or it's not. If it is consistent, then (2) might be true but it can't undermine the philosophical theory. If it is not consistent with the correct philosophical theory of well-being, then it's the <u>empirical</u> claim that is in jeopardy. To see this, suppose that the psychologists who performed the study presupposed a mistaken view about the nature of well-being: they found that meditation tends to bring about positive hedonic states of the sort the hedonist would identify with well-being, but hedonism, as a philosophical theory of well-being, is false. In this situation, we would conclude that even though the study has shown <u>something</u> (e.g., that meditation tends to bring about positive hedonic states), it hasn't necessarily shown that meditation tends to foster <u>well-being</u>. And so we see that the philosophical theory of well-being is invulnerable to psychological findings (Isolation) and psychological findings purporting to be about well-being are vulnerable to philosophical theorizing (Vulnerability).

The traditional approach to the study of well-being has all the benefits of incumbency. Philosophers understand it and are comfortable with it. But it makes at least three substantive assumptions that are open to question.

1. Everyone's "preanalytic convictions" about the nature of well-being are subserved by (roughly) the same conception of well-being.
2. The conception of well-being underlying our "preanalytic convictions" about the nature of well-being can be accounted for in terms of singly necessary and jointly sufficient conditions.
3. Everyone's "preanalytic convictions" about the nature of well-being are epistemically privileged – they are so worthy of belief that it is appropriate that they form the primary evidential base for a theory about the nature of well-being.

None of these assumptions is obvious. As to the first, recent empirical work in psychology and philosophy suggests that people in different cultures and socioeconomic groups diverge, sometimes dramatically, in some of their philosophically significant judgments (citations). Focusing on the topic at hand, we find a fairly wide range of views about the nature of well-being even among Western philosophers. And sometimes these differences lead to disagreements about whether some intuitions (or "preanalytic convictions") about well-being really stand up to scrutiny (e.g., Crisp 2006, yy). It is not difficult to imagine systematic cross-cultural differences along the same lines – for example, one society made up primarily of hedonists and another made up primarily of informed desire theorists. In each society, people are unmoved by the standard objections to their culture's theory; so folks in the hedonist society find Nozick's Experience Machine objection completely unconvincing – as long as the subjective lives of Nozick's subjects are identical, it seems perfectly intuitive to folks in that society that the person who is having the real experiences and the person who is attached to the Experience Machine have the exact same level of well-being. I am not arguing that there really is this sort of cross-cultural diversity in people's conceptions of well-being; I am merely pointing out that it is possible for different people, and so different cultures, to have quite different views about the nature of well-being. What's more, insofar as we think that philosophers who embrace different explicit views about the nature of well-being can still talk about well-being, it is reasonable to assume that non-philosophers whose preanalytic convictions are subserved by different implicit conceptions of well-being can still talk about well-being. As to the second assumption, many philosophers (notably Wittgenstein and Searle) and many psychological studies suggest we do not represent many of the categories we employ in terms of singly necessary and jointly sufficient conditions (citations). For example, some categories might be represented by typical examples, or exemplars, and a person might judge whether a target item is a member of the category in terms of how similar the target is to the relevant exemplars. It is possible that our commonsense judgments about S's well-being are the result of how similar we deem S to be to our relevant exemplars of well-being. If something like this were true, it would be unlikely that we could account for those judgments in terms of a theory framed in terms of singly necessary and jointly sufficient conditions. As for the third assumption, let's assume that well-being is a real condition in the world. It is not uncommon for many of our reflective pretheoretical judgments about real conditions in the world to be mistaken. I'm not insisting upon a deep skepticism about our commonsense judgments about well-being. I am simply suggesting that

perhaps a bit of epistemic modesty is called for when it comes to our commonsense judgments about the nature of real conditions in the world. It is certainly possible that our commonsense judgments about well-being are so spectacularly accurate that they deserve to be taken as the primary evidence for our theories about the nature of well-being. But is it really a good idea for everyone in philosophy working on well-being to rely on this assumption? After all, it might be that our commonsense judgments about the nature of well-being are just fairly accurate – they're generally true but systematically mistaken in a few ways. We can worry about the extreme epistemic optimism implicit in the third assumption without thereby committing ourselves to an extreme pessimism about our commonsense judgments about the nature of well-being.

Proponents of the traditional approach to the study of well-being might offer a number of arguments for thinking that our judgments about the nature of well-being are so worthy of belief that we <u>must</u> accept them as the primary evidential base for a theory about the nature of well-being. Some might argue that we employ a faculty of rational reflection that puts us in a position to know the essence or nature of well-being. Others might argue that we are actually investigating what we <u>mean</u> by the expression 'well-being' and the reliability of this investigation does not depend on whether the empirical propositions expressed by the sentences in which the expression 'well-being' occurs are true. Or they might argue we are trying to capture the content of the <u>concept</u> of well-being, where we can suss that content even if it is sometimes embedded in false empirical beliefs (or believed propositions). Each of these possibilities raises further worries. If we are using some faculty of rational reflection, what is this faculty? And why should we think that it is <u>so</u> reliable that it is reasonable for our pretheoretic judgments to make up the primary substantive evidence for a theory of well-being? And if we are investigating the meaning of 'well-being' or the concept of well-being, why should we think that this meaning or concept accurately reflects the true nature of well-being? There may be powerful answers to these concerns. And we could explore these various rabbit holes for a long time. But it is worth noting that besides making non-obvious, substantive assumptions, the traditional approach has not generated anything close to consensus in the philosophical community about the nature of well-being. Maybe it just needs more time. But why put all our eggs in that basket? Given this situation, it is reasonable and appropriate to consider an alternative approach to investigating the nature of well-being, one based on more modest assumptions.

3. The first step toward a naturalistic approach to the study of well-being

The naturalistic approach to the study of well-being I will defend begins with an important and widely accepted feature of the work of Kripke (1972) and Putnam (1975) on natural kinds: We can talk about something even when we're quite ignorant or mistaken about its real nature. People who had deeply mistaken or incomplete views about the nature of water, electricity, electrons, atoms, planets, stars, meteors, asteroids, combustion, disease, and light (this list could go on and on) nonetheless were able to talk about (refer to) those things. Of course, the Kripke-Putnam semantic machinery consists of a lot more than this modest insight – but the naturalistic approach to well-being proposed here is not committed to any specific views about the nature of natural kinds or how natural kind terms refer. The first step toward a naturalistic approach to the study of well-being is to recognize that we can talk about well-being even if we're quite ignorant or mistaken about the real nature of well-being.

The assumption that we can talk about well-being even though our pretheoretic views about its nature can be mistaken undercuts the traditional approach to the study of well-being. It implies that we cannot <u>begin</u> our investigation into the nature of well-being by making the bold assumptions implicit in the traditional approach. Consider again the three substantive assumptions of the traditional approach broached in the previous section. If we assume that we can talk about well-being even though we are somewhat mistaken about its nature, then we are no longer committed to any view about those three theses. They are genuinely open questions.

1. The naturalistic approach leaves it an open question whether everyone's pretheoretical convictions about well-being presuppose roughly the same general account of well-being. Suppose different people's pretheoretical convictions are subserved by different general views about well-being. So George's convictions are grounded by a hedonistic view whereas Sally's are grounded by an informed desire view. Their commonsense judgments about well-being will overlap considerably. For example, they will agree that Susan (introduced in chapter 1) has a high degree of well-being and they will largely agree about the factors that are involved with her well-being – her positive feelings, successful relationships, professional success, etc. Of course, their explanations for why these factors contribute to their well-being will be different - George thinks they contribute to Susan's well-being because they bring about happiness or pleasure, while Sally thinks they contribute to Susan's well-being because they satisfy her informed desires. So while there will surely be some differences in their judgments, as long as the full range of their convictions about well-being do a reasonably good job of tracking well-being, it makes sense to suppose that they're both talking about well-being.

2. The naturalistic approach leaves it an open question whether the general account of well-being that grounds a person's pretheoretical convictions about well-being can be framed in terms of singly necessary and jointly sufficient conditions. In fact, many people's commonsense conception of well-being might be quite confused, vague or ambiguous. As long as their judgments about well-being track facts about well-being reasonably well, they're talking about well-being.

3. The naturalistic approach leaves it an open question whether our pretheoretical convictions about the nature of well-being are so coherent and so accurate that we can extract the true theory about the nature of well-being from them. They <u>might</u> be highly accurate; but they might be quite mistaken or confused. The naturalistic approach does assume that well-being is not a failed posit, like caloric or aether. As long as our judgments about well-being aren't so deeply confused that they fail to reasonably track anything, it makes sense to suppose those judgments are about well-being.

The traditional approach assumes that our reflective, pretheoretical judgments about well-being are so coherent and so accurate that we can extract the true theory about the nature of well-being from them. Perhaps this bold, optimistic assumption is true. But deep confidence about our commonsense judgments about well-being should be the <u>result</u> of our investigation into the nature of well-being, not the <u>starting point</u> of that investigation. We need an approach to the study of well-being that doesn't begin by assuming the immaculate perspicacity of our commonsense judgments. And that's what we get from the first step toward the naturalistic approach. But where do we go next?

4. Roadblock: Is there really an alternative to the traditional approach?

Suppose we take the first step toward a naturalistic approach and we reject as premature the optimistic assumptions behind the traditional approach to the study of well-being. So we do not <u>begin</u> our study with the assumption that our commonsense judgments about well-being are so epistemically special that accounting for those judgments is the "primary test" for a theory of well-being. A problem immediately presents itself: If we are not going to rely on our commonsense judgments about well-being in constructing our theory, what <u>are</u> we going to rely on? An obvious answer, and one that we would expect from a naturalist, is: Science. But prima facie, this suggestion is deeply unpromising. No empirical study can confirm or disconfirm hedonism – or any philosophical theory about the nature of well-being. As Daniel Haybron notes with respect to theories of happiness: "What kind of empirical study could possibly tell us which account [of happiness] is correct? One might as well try performing an experiment to determine whether water is H_2O or a kind of bicycle" (2003, 312). But if we can't rely on our commonsense judgments and we can't rely on science, what's left? It appears we have hit a roadblock. If so, maybe we should go back to the traditional approach. A flawed approach is better than nothing.

To see whether we can get by this roadblock, let's explore in more detail the claim that science can't help us construct a theory of well-being. We can very roughly divide the psychological literature on well-being into two kinds of research:

(a) theoretical speculations about the nature of well-being, and
(b) empirical studies that identify factors that correlate with or causally interact with some measurable condition that is taken to be (or to accurately measure) well-being.

It seems that neither of these sources of information can provide us with a <u>distinctive</u> method for investigating well-being – a method that is different from the traditional one. With respect to (a), psychologists don't speak with one voice about the nature of well-being (citation). But even if they did, it would be naïve to uncritically accept the speculations of psychologists about the nature of well-being. So how are these proposals from psychologists about the nature of well-being to be evaluated? This is exactly the question we started with! It might or might not behoove philosophers to pay more attention to psychologists' speculations about the nature of well-being. But it would seem that we must evaluate their proposals according to the standards implicit in the traditional philosophical approach to the study of well-being - by seeing how they account for our "preanalytic convictions" about well-being. And so (a) drives us right back to the traditional method.

What about (b) – the empirical studies that tell us what conditions tend to foster, impede or correlate with well-being? The traditional approach is committed to Isolation, which says that the primary evidence for or against a theory of well-being comes from our pretheoretical judgments, and empirical findings about the causes, effects or correlates of well-being are not relevant evidence for or against a theory about the nature of well-being. So (b) is irrelevant. Let's consider the traditionalist's Isolation argument as it might apply to a real study. There are many instruments psychologists use to measure well-being, but here is a common one, the five-item Satisfaction with Life Scale, or SWLS (Diener 19xx).

Below are five statements that you may agree or disagree with. Using the 1 - 7 scale below, indicate your agreement with each item by placing the appropriate number on the line preceding that item. Please be open and honest in your responding.

7 - Strongly agree
6 - Agree
5 - Slightly agree
4 - Neither agree nor disagree
3 - Slightly disagree
2 - Disagree
1 - Strongly disagree

____ In most ways my life is close to my ideal.
____ The conditions of my life are excellent.
____ I am satisfied with my life.
____ So far I have gotten the important things I want in life.
____ If I could live my life over, I would change almost nothing.

The SWLS scores are interpreted as follows: 31-35 is extremely satisfied; 26-30 is satisfied; 21-25 is slightly satisfied; 20 is neutral; 15-19 is slightly dissatisfied, 10-14 is dissatisfied; and 5-9 is extremely dissatisfied.

We needn't get bogged down (yet!) in the merits of this particular way of measuring well-being. The general point to make here is that in order for any empirical study to make a discovery about well-being, it must employ an instrument that measures well-being (or a proxy for well-being) reasonably accurately. Take any instrument used by psychologists for this purpose, like the SWLS. Why suppose that it tells us anything about well-being? As we've already seen, the traditionalist offers the following argument schema for why any such empirical finding is irrelevant to the philosophical project of accounting for the nature of well-being: An empirical study can inform us that (say) reported self-esteem or income is (or is not) strongly correlated with SWLS results, or that beginning a meditation regime tends to raise SWLS scores. But whether such results tell us anything about well-being depends on whether SWLS plausibly measures, or is a proxy for, well-being. But this is something we can know only if we presuppose a philosophical theory about the nature of well-being.

We can see this argument schema play out in Valerie Tiberius's discussion of some fascinating evidence that suggests that well-being has somewhat different profiles in different cultures (2004). For example, self-esteem correlates better with well-being in individualist cultures than in collectivist cultures (Diener and Diener 1995). Tiberius wonders whether this kind of cross-cultural diversity, if it were ubiquitous, might undermine the philosophical project of providing a unified, universal theory of the nature of well-being. She argues it does not. In order for an empirical study to be about well-being, it "must presuppose an account of the nature of well-being" (Tiberius 2006, 295). And so a formal, philosophical theory of well-being is prior to and presupposed by empirical findings about well-being. "[I]t seems that some such [formal] account is required if psychological or philosophical research is to be done on cultural or

individual differences in the constituents of well-being" (307). This is because "psychologists investigating cross-cultural variance in well-being have to start with some notion of well-being that is common to both cultures… Without this assumption, it would have been impossible to say that these differences between cultures are differences that have to do with well-being" (295).

I want to argue that we can bypass this roadblock. We can learn about the nature of well-being from (b), from empirical studies that identify factors that correlate with or causally interact with some measurable condition that is taken to be (or to accurately measure) well-being. But we need to proceed with a bit of care. Let's first explore the traditionalist's contention that empirical studies always presuppose some "notion" or "general account" of well-being. The plausibility of this claim depends on how we understand "notion" or "general account." It is certainly <u>not</u> true that scientists who study well-being <u>must</u> presuppose some well-developed philosophical theory of well-being (like hedonism) - although of course they might. And it doesn't follow that all scientists who use the same instrument for measuring well-being <u>must</u> therefore share the same understanding of well-being. Consider the SWLS. As we shall see, there are reasons to worry about this and other self-report measures of well-being (Schwartz & Strack 1999). But putting aside these concerns, a proponent of any plausible theory of well-being can agree that the SWLS is not an outrageous way to measure well-being, on a par with measuring well-being in terms of eye color. Of course, SWLS might <u>fit better</u> with some theories than others. And so someone who thinks people's life satisfaction judgments are essential to well-being (e.g., Sumner's authentic happiness theory) will probably argue that SWLS is a <u>more accurate</u> measure of well-being than someone who thinks that well-being involves people's pains and pleasures (e.g., a hedonist). But there is no theoretical reason for why both the hedonist and the informed desire theorist cannot agree that SWLS measures well-being reasonably accurately. What is surely true is that in order to study and learn about well-being, scientists and non-scientists alike must have <u>some</u> understanding of well-being. But from this pedestrian truth, not much follows about how to properly investigate the nature of well-being.

Proponents of the traditional approach might reply that at this point, they have lost the battle but won the war. If empirical studies aren't committed to a single, specific, well-developed theory of well-being, then how does the naturalist propose to extract a specific, well-developed theory of well-being from them? This is the right question to ask! But to think that this question raises a serious problem, one must assume that the naturalist proposes to simply "read off" a theory of well-being that lies implicit in particular empirical studies or methods (e.g., measuring well-being with the SWLS). But this would be a rather obtuse, flat-footed naturalism. Consider an analogy. In order to study and learn about the evolutionary history of organisms, biologists and non-biologists alike must have some understanding of fitness. But the sophisticated philosopher of biology knows she can't simply "read off" the correct account of fitness from scientific studies or from scientists' understanding of it. Not every scientist who talks about fitness understands it in precisely the same way; and even if they did, that understanding might be in some ways flawed. What's more, the mere fact that in order to talk and learn about fitness, everyone must have some understanding - though not necessarily precisely the same understanding - of fitness will not incline the sophisticated philosopher of biology to throw up her hands at the prospect of providing an empirically informed, unified account of fitness. By analogy, the fact that in order to study and talk about well-being we must have some understanding - though not necessarily precisely the same understanding - of well-being does

not, by itself, undermine the possibility of developing an empirically informed, unified account of well-being. It is a mistake to suppose that the proper way to extract a theory of well-being from empirical studies is to somehow straightforwardly "read off" a theory from such studies.

On the view I am pressing, the right response to the cultural diversity evidence is not to insist that philosophy has nothing to learn from psychology about the nature of well-being - that no matter what psychologists tell us about well-being in different cultures, we know there must be a unified conception of well-being that psychologists start with in their studies. Instead, the right response is one that keeps the question open: The philosophical project of finding a universal account of the nature of well-being is perfectly consistent with cross-cultural diversity. They are consistent because the underlying nature of well-being might be cross-culturally uniform even if the factors that promote or inhibit well-being are different in different cultures. This is a compatibilist reply. It says only that the empirical cross-cultural evidence does not, by itself, rule out the possibility of a universal theory of well-being. But we have no guarantee that we will find such a theory, either. The question of whether we can find a theory of well-being that applies to all human beings is an open one. And given the current state of our knowledge, that seems like the right answer.

We began this section facing a roadblock: The traditionalist had an argument for why empirical results purporting to be about well-being could not serve as evidence for or against a theory about the nature of well-being. In reply, I have argued that this is a roadblock only if we assume an addled sort of naturalism - a naturalism that supposes we can unproblematically find the correct theory of well-being lying implicit in a bunch of empirical studies. We can bypass the roadblock by rejecting this flat-footed conception of naturalism, but the way forward remains a mystery. I'm still on the hook to explain what resources are available to the naturalist who wants to develop a theory of well-being that is informed by the empirical literature. How might a sophisticated naturalist proceed in her study of well-being? It is to this question we now turn.

5. A positive naturalistic approach

The naturalist approach I will press assumes that well-being is a state, W, that scientists have been learning about and that laypeople talk about, perhaps imperfectly, when we talk about well-being. I propose that W is well-being if:

1. W is a causally stable condition that can be identified and characterized by standard empirical methods (the causation criterion);
2. W helps to organize, unify and explain significant features of the empirical literature on well-being (the coherence criterion); and
3. The assumption that W is well-being (and thus that 'well-being' refers to W) makes most of our commonsense judgments about well-being true or approximately true (the commonsense criterion).

Let's suppose some state, W, meets the causation, coherence and commonsense criteria. By abduction, it follows that W is well-being. In other words, the best explanation for why W meets these three criteria is that W is well-being. In order to motivate this approach, I want to say a bit more about each criterion.

The first criterion holds that well-being is an empirically stable condition. I will argue that this stability is explained by the fact that well-being consists of various self-maintaining and self-reinforcing causal processes. Well-being is a homeostatic property cluster: it is made up of a family of properties that tend to co-occur because "[e]ither the presence of some of the properties… tends to favor the presence of the others, or there are underlying mechanisms or processes which tend to maintain the presence of the" property cluster (Boyd 1989, 16).[1] We will explore this and its implications in more detail in chapter n.

The empirical literature on well-being is massive and unwieldy. The second criterion holds that well-being is a theoretically important posit – an accurate understanding of well-being should help to organize, unify and make sense of this literature. Consider one way in which a proper understanding of well-being might bring some theoretical order and regimentation. The psychological literature on well-being consists of a motley of different approaches and involve lots of different instruments putatively measuring well-being or its components. Many of these approaches (e.g., emotion-focused, interpersonal, coping) seem like intuitively plausible entrées into understanding well-being; and many have yielded interesting empirical results. But how are we to make sense of this multiplicity of approaches and measures? Are some more central to the study of well-being than others? Or are they all equally valid ways to understand human well-being? A proper understanding of the nature of well-being will help us to give compelling answers to these questions. Chapter 4 will spell out the case for thinking that positive causal networks satisfy the first two conditions, the causation and coherence conditions.

It is not enough to show that W is an empirically stable condition and a posit that is important for properly understanding a large and important empirical literature. There are many such posits that aren't well-being - planets, stars and gold, for example, meet those two conditions but aren't well-being. We need some reason to believe that the condition under consideration is well-being. Proponents of the traditional approach are right to note that well-being is not in the first instance a scientific notion. As Sumner states, "Because the notion of welfare [well-being] already has a vernacular currency it is not available as a term of art, to be defined in whatever way will best suit some favoured theoretical needs…" (1996, 10). Daniel Haybron makes the same point with respect to happiness: "happiness is… a folk notion; 'happiness' is not a technical term for theorists to use as they please" (2003, 312). While 'well-being' was part of our common vernacular long before psychologists began systematically studying it, the same is true of many posits whose natures have been unlocked empirically by science – water, gold, human. So the mere fact that well-being has "a vernacular currency" implies very little about how to properly study it. In particular, as I have argued, we need not be particularly deferential to our commonsense ideas about well-being. But this makes the problem particularly acute: If W is not faithful to our pretheoretical, commonsense ideas about well-being, then why think W is well-being? This is where the third criterion enters.

An important assumption behind the approach on offer is that we can talk about well-being even though our views about the nature of well-being are incomplete or mistaken. This assumption involves a sense-reference distinction. According to the Kripke-Putnam view of

---

[1] Technically, HPCs are made up of entities or events that possess or instantiate properties, not the properties themselves.

natural kinds, the expression 'water' refers to H2O even if the descriptions speakers typically associate with water (the sense of 'water') are mistaken. Just as 'water' refers to H2O for those whose concept of water is innocent of modern chemistry, 'well-being' refers to a stable condition, W, even for those whose concept of well-being is innocent of modern psychology. On this view, people who hold quite different views about the nature of water can still talk about (refer to) water as long as most of what they say about water is true (or approximately true) about water. (The same point holds for light, planets, stars, etc.) So to make the case that W is well-being, it is not necessary to show that W accords in every detail with our commonsense judgments about the nature of well-being. Rather, what we need to do is show that assuming that 'well-being' refers to W makes most of what we say about well-being true or approximately true. To make this case, we're going to need a fund of clearly stated and intuitively plausible commonsense judgments about well-being. We can then test whether the hypothesis that the candidate condition is well-being can account for a significant number of these judgments. But where can we find such a fund of judgments? From the philosophical literature on well-being, of course! My plan is to argue that the assumption that well-being is W makes approximately true much of what philosophers say about well-being – both when proposing and criticizing theories of well-being.

This approach is unconventional. Because our understanding of well-being might be imperfect, the goal will not be to show that identifying well-being with W captures our commonsense understanding of well-being better than competing theories (e.g., hedonism or informed desire theory). Rather, the core argument will take the form of an inference to the best explanation: The assumption that W is the referent of 'well-being' makes sense of much of what philosophers say about well-being – both when proposing and criticizing theories of well-being. The best explanation for this "fit" is that 'well-being' refers to W. Chapter 5 will make the case for thinking that positive causal networks satisfy the third and final condition, the commonsense condition.