

Measuring “The Good Life”: Alternative Metrics for National Wellbeing

Heather Howard 00:00

Welcome. Today's episode is all about "the good life." We've heard the phrase countless times. But what does living the good life really mean? And how can we measure our success? Those questions are the subject of great debate among researchers, policymakers, and other stakeholders who strive to understand the many dimensions of human welfare, and how each of those dimensions contributes to a nation's wellbeing.

Within the realm of economics, Gross Domestic Product, or GDP, has been the key indicator for evaluating a country's performance and prosperity. But there's a growing movement to look beyond that metric. GDP only accounts for goods and services, it fails to consider health, access to education, happiness, and other things that people care about.

That's why many economists argue the need for a better yardstick and a more complete picture of national wellbeing. Among those economists is Professor Ori Heffetz from Cornell University and the Hebrew University of Jerusalem. A Princeton alum and visiting research scholar, Professor Heffetz is here today with Dr. Yanchun Zhang, chief statistician at the United Nations Development Program. With more than 20 years of quantitative research experience, Yanchun is a proponent of the Human Development Index, an alternative approach to assessing human welfare and rating a country's wellbeing. On today's show, the three of us will explore that index, along with other strategies for creating a more accurate, comprehensive, and useful system for measuring "the good life" in societies around the world. Ori, Yanchun, welcome!

Ori Heffetz

Thank you. Great to be here.

Yanchun Zhang 02:14

Thank you, Heather.

Heather Howard

Ori, let's start with you. Tell me about the GDP and why you think it's a flawed measure of wellbeing.

Ori Heffetz 02:25

GDP was never meant to be a measure of wellbeing. In fact, it is something that was conceived between the world wars and then got a boost during World War II. But it was an attempt to measure the productive capacity of national economies in the context of war, and later in the context of peace -- to count everything that an economy can produce. In the long run, everything we can produce is also everything we can consume. So it's a good measure of economic living standards maybe, but there is

more to life than economic living standards, right? So people want to live according to their personal values. And people want to have freedoms, all sorts of freedoms, and they want their political voice to be heard. And they want to live in societies they perceive as just, and they want to have good quality of family and social relationships, and they want to have a pollution-free environment, and they want to have a lot of things that are not counted in this measure.

Heather Howard 03:38

Yanchun, do you agree that GDP is a flawed measure? And does it get used in ways that it shouldn't be?

Yanchun Zhang

I agree with what Ori said. In particular, in the historical context, it was developed, as already mentioned, in 1937, by Simon Kuznets [a U.S. economist], and then got a lot of attention after World War II, especially after the Bretton Woods Institution adopted it as a basic tool to evaluate the size of the economy. But since the 1970s, there has been so much criticism. Even in 1972, Bhutan started proposing "Gross National Happiness" as the goal instead of GDP as a goal for national development. So, yes, there have been a lot of misses from GDP.

In addition to what Ori mentioned, common criticisms include that GDP does not account for non-market transactions, like unpaid care. Also, inequalities are not accounted for. The GDP measures gross flows, so the depletion of natural capital is not taken into account. All of these issues are common criticisms that, increasingly, policymakers feel is not an ideal measure for national wellbeing or individual wellbeing.

Heather Howard

So let's play that out. If you were using GDP as your only measure to see national wellbeing, how would different countries look during COVID? And how might the GDP, as a flawed measure, skew what you find? Can you give me examples -- what it might miss, what it might tell us, and how we might draw the wrong inferences.

Ori Heffetz 05:35

COVID is a great example. Early in the pandemic, people talked about a trade-off. On the one hand, you could keep the economy open, which would minimize the damage to GDP, but that could get more people infected. So there was perceived trade-off. If you shut down the economy, you may save lives, but you may kill some economic activity. At the end of the day, I'm not sure there was a trade-off because saving lives or stopping the COVID wave is also good for the economy in the long run. But that is a great example of how, if you are only trying to maximize GDP in a time like the COVID pandemic, that strategy would be a ridiculous and embarrassingly indefensible policy goal because people's lives are on the line.

Heather Howard 06:42

What's another example? I read in, in your work, about traffic as example. Traffic may show up as actually adding to the GDP, but I think many of us don't want to be in traffic. And so we wouldn't consider it a measure of wellbeing. Can you speak to that, Yanchun?

Yanchun Zhang

Sure. If I may, I'd like to quote our Founding Director of the Human Development Report office, Mahbub ul Haq, who once said that any measure that places the value of a gun 100 times more than a bottle of milk should be seriously debated as a measure for human progress. I think linking to COVID, you could also say that, right? If you look at GDP only, it doesn't capture the devastating impact of this global poverty crisis on the humanity, not at all. It's just a one thing. There are also a lot of other indicators that may show big fluctuations, post-COVID. But it doesn't really reflect the human loss.

Heather Howard

And why should it matter that we're not measuring it? There's the maxim that says "what gets measured, gets managed," or what you measure is what you get. Is that why it matters? You're in a policymaking role at UNDP, so how do you see this playing out? And why should people care?

Yanchun Zhang

What we measure decides what we do. It also gives us a precise, or pretty precise, sense of the current situation. That's why it's very important to have not only GDP, but also other measures. We do see the value of GDP, but it's definitely not enough. As a wellbeing indicator, it's flawed. So do you worry that policymakers, by depending on a flawed measure, are drawing those wrong influence inferences and that poor policymaking results because of it? Should we be looking downstream to the impact of having this flawed measure?

Ori Heffetz

I think so. If we want our policymakers to conduct data-based policy -- not to dream up policies or intuit policies, but to look at research and to locate estimated impact of different policies and to act by that. We are moving more and more in this direction. There is more and more data available, for better and for worse, etc. So if we take this approach of data-based policymaking, then the data you look at will determine the policy. If the only metric that you trust, and that exists for almost 200 countries for many years, is GDP, you will end up looking a lot more at that than at other metrics that may not be developed yet, or don't yet exist for all the countries, or maybe don't even exist at all. So I think a data collection effort is the first step towards data-based policy that will take into account more than one metric.

Heather Howard 10:30

Let's talk about that. What are alternatives to the GDP? Yanchun, you've been working on the Human Development Index. Can you tell us how that came about and how you think it is a better alternative to GDP and takes a more multi-dimensional look?

Yanchun Zhang

Sure. The Human Development Index (HDI) was launched in 1990, when the first Human Development Report was published. HDI is never a stand-alone index; it's part of the Human Development Report. The whole report series and index and other composite indices we publish are based on the human demand approach, which is also closely linked to the capabilities approach to wellbeing and human development.

Heather Howard

Before we go on, can you say more about that capabilities approach? Because that was a new line of thinking in economic research, right?

Yanchun Zhang

In simple terms, the capability approach to wellbeing, for example, argues that an individual's wellbeing depends on whether this individual has the ability to exercise freedoms and choices of "beings" and "doings." "Beings" are what you want to be. For example, you want to be healthy, you want to be educated...these are the beings. "Doings" are the things you want to do. For example, you want to have a decent standard of living, or you want to have a good relationship with your families and colleagues. So, at a very basic level, the core capabilities include the ability to read or write, to be healthy, and to have a decent standard of living. These capabilities all contribute to the construction of HDI.

Heather Howard

Those are lofty goals. I was on your website. Let me just quote how it's described on the UNDP website, "HDI was created to emphasize that people and their capabilities should be the ultimate criteria for assessing the development of a country, not economic growth alone." This follows the conversation we just had about the flaws with GDP. So how do you translate those into the dimensions of wellbeing that you're tracking, and the relative importance of each. What goes on behind the curtain in the calculations here?

Yanchun Zhang

HDI is a multi-dimensional capabilities index. It has three dimensions: health, education, and standard of living. It's a very simple composite index. For the health dimension, we use life expectancy at birth as the indicator. Then, for the education dimension, we'll look at two indicators: expected years of schooling and mean of years of schooling. For standard of living, we look at gross national income per capita. So these four indicators go into the construction of HDI. We calculate HDI every year for 192 countries and territories. Then we also update the time series of HDI from 1990, the year the first report was published until the latest year. So now we have a pretty long time series for countries to use this index to evaluate its own progress over time. And also you can use this consistent series to compare your performance with your neighbors' performance. It's a nice nudging tool for countries to think about which dimensions they can improve on.

Heather Howard

Ori, as researcher, how do you use the HDI and how do you react to the efforts to create a more effective measure? What do you think of HCI?

Ori Heffetz 14:42

First of all, the HDI is a great success and an example that we often talk about when we say that the role of GDP is not predetermined and we can move beyond it. The HDI successfully moved beyond a uni-dimensional metric, income alone, and put in two additional dimensions of wellbeing, arguably two of the most important: health and education, or you could call it access to knowledge. That's a great leap forward. Now we have a metric that has three important dimensions in it rather than one. What we show in our research, or study, or try to convince policymakers and statistical national agencies, is that three is a great start, but why stop at three? There are other important dimensions of wellbeing that go beyond income, health and education. Let's add them, just like we already started adding things.

Heather Howard 16:09

So this is an iterative process, as far as you're concerned. You think you can build on it. I want to hear about your ideas for building on it. But before that, I'd like to ask you how the HDI might be helpful in comparing high-, low-, and middle-income countries. How might the reliance of just looking at wealth skew comparisons between those countries? Is the HDI helpful in, for example, studying questions about equity between high-, middle- and low-income countries?

Yanchun Zhang

Of course. A lot of scholars, policymakers and general readers use our index when they compare HDI values and ranks versus GDP values or ranks. Some high-income countries are not ranking so high in our index because of the other two dimensions. But some countries with relatively low GNI per capita can have a quite nice rank, according to HDI, because they have relatively better performance in the health and education dimensions.

Heather Howard

So the United States, for example, looks quite different under the HDI than we do just on a pure GDP measure, right?

Yanchun Zhang

Yes. In this year's report, we published the latest HDI. We then compared that to the last report, and also looked at the time series of U.S. HDI values or ranks, and found that the country has been declining. After COVID, the U.S. dropped out of the "Top 20" country list.

Heather Howard

Thank you. I know people in the United States don't like to hear that, but it's an important conversation to have.

Yanchun Zhang

Yes. One of the dimensions to health -- life expectancy at birth -- took a bit hit in the U.S. after COVID.

Heather Howard

But even before COVID, we saw a decline in life expectancy in the U.S. Some Princeton professors actually coined the term "deaths of despair." Anne Case and Angus Deaton have done seminal work on that front, and I think you've worked with them, Ori, right? There's a Princeton connection everywhere.

Ori Heffetz

Yes. So actually, Anne Case was the chair of my Ph.D. thesis committee, and Angus Deaton was a member of the committee. So the seeds were planted for the stuff I'm working on today back when I was younger grad student at Princeton.

Heather Howard

Ori, you mentioned that the HDI is really a transformational leap from GDP -- that we now have significant data and longitudinal data, which is great, but you want to add to that. What are your ideas for expanding on these measurements? What else would you like to add?

Ori Heffetz

There are things that we find in our research as consistently important -- if your criterion for importance is what people tell you in surveys that we conduct. Beyond physical health, they also persistently talk about mental health. People talk a lot about family wellbeing. People talk a lot about security or sense of security of all sorts. They often talk about financial security, a general sense of security about the future. People talk about freedoms a lot. I mentioned that before... freedom of speech, freedom of mobility. For example, freedom of mobility was curtailed overnight during COVID, in a lot of the world. One day we just couldn't go everywhere that we wanted to. I'm talking about the findings of work that we have done over the past 10 years, way before COVID. These are things that are important to people and that the current metrics don't take into account. The GDP definitely doesn't consider those things, but also the HDI doesn't yet take them all into account either. We would like to see governments collecting this data. The UNDP cannot publish once a year a metric that relies on dimensions that not all of the member countries actually measure in a reliable and comparable way. So, as the first step, let's start collecting data on more dimensions that we think are important. And then the second stage would be to collect data on the weight, or relative importance, of each dimension to people. Then we will be able to aggregate all of these different dimensions into one index.

Heather Howard 20:48

Yanchun, can you talk about that. Data collection is not easy, right? But do you have thoughts about how you can get to that future in terms of scaling up data collection?

Yanchun Zhang

I couldn't agree with Ori more, but we don't collect data. We are users of data in our office, because we rely on other UN agencies. For example, for educational data we rely on UNESCO Institute of Statistics. Then for life expectancy data, we use official data from the UN Population Division. This data is collected by these custodian agencies, harmonized, standardized, then published. So to ensure the transparency of our index calculation, and also to ensure the comparability of the index, we use the official data from the UN agencies and other international organizations. But it doesn't mean that we are not involved in efforts to obtain more timely data collection. This is super important. After COVID, it's clearer than ever that we need more timely data to evaluate the impact of COVID, to be able to distribute vaccines to the areas where they're most needed.

Heather Howard

That's interesting. So COVID has highlighted further the need for timely data - that if you're acting on old data, you're not being as effective as you can be in targeting outreach. Ori, you've written about time lags in data collection, too, and how frustrating that can be.

Ori Heffetz 22:33

Yeah, so that is a trade-off, a painful tradeoff, in economics and other fields in economics. We want to make policy. I keep going back to this phrase of data-based policymaking. We want our policymakers to rely on data. Certainly during a pandemic that moves very fast, we want policy to react fast on the most recent data. But we also want policymakers to base policy on the most reliable, harmonized, standardized, transparently collected data. And that's always a trade-off -- especially when something hits you overnight, like the COVID crisis, and you're realizing there's a lot of data you would like to have right now that you have never even thought about collecting. So you start very quickly collecting it, somehow. You understand that it's not going to be perfect, but you'll have something to act on tomorrow morning. That's exactly when these trade-offs are the most painful. Are you going to wait, and verify and clean the data? Or are you going to do something quick and dirty, which may save more lives?

Heather Howard 23:48

That's a constant challenge for policymakers, right? Do you wait for the perfect data, or use data when it's "good enough"?

Ori Heffetz

Exactly. I think, in "normal times" -- and who knows anymore what "normal times" are -- we don't need the data for tomorrow morning. So we do have a little bit of luxury to think about and conduct research on the missing dimensions. What are the things that are important to people's wellbeing that we are not currently collecting on? Let's start working on it now, because it's going to take a long time to develop the metrics, to standardize them, to convince more countries to measure them. Let's start doing it now. So that next time we need more data than income data, we have it.

Yanchun Zhang 24:39

Exactly. We need to have this ex-ante, not ex-post. This time there's a global public health crisis that affects so many people, so many countries, and it got a lot of attention. But we have crises all the time. Previously we had the Ebola crisis, then we have regional conflicts, we have commodity price volatility that affects exporting developing countries. The impacts of all these shocks need to be examined, and data collection is crucial.

Heather Howard

Ori, you mentioned trade-offs in the context of data collection. But there's another kind of trade-off that you want to explore, which is how individuals weight their preferences. Can you say more about that and your ideas for addressing it?

Ori Heffetz

Yes. So first, just to clarify, these are not my ideas alone. I've been part of a great team of researchers that include Dan Benjamin, now with UCLA, and Miles Kimball, now at Colorado, and Kristen Cooper,

now at Gordon College. If we want to develop an index that is no longer uni-dimensional, no longer relies on only one thing, for example, income, then we need to do two things. First, we need to collect data on more dimensions. And we have talked about that in the last few minutes. Then, at some point, once we have data on more dimensions, how do we combine them into an index? Which of these dimensions get more weight, which get less weight? This is a normative decision that could affect the outcome, of course. How much weight you give different things could affect the outcome. So it's a normative question that society needs to ask. And there is a lot of discussion among academics and among philosophers and among policy makers and in the public media, about the things that are important to society. What are they?

Our idea is to conduct surveys among the relevant population, and to use standard tools that economists often use to price things or to get at the trade-offs between different things. We can use these tools to estimate, for example, how much people value time. This is something that economists often do, right...they put a number and dollar value on time. We're saying let's use exactly the same techniques, but adapt them to price, health, physical health versus mental health, versus security about the future, versus income, versus access to knowledge and education. We use different types of surveys, where people face different scenarios and make hypothetical choices between two situations. For example, they might choose between a situation where you have slightly more health, let's say, and a situation where you have slightly more income. We use these choices that people make in the surveys to estimate weights.

Heather Howard 28:35

As I read in one of your papers, you're testing personal preferences and then aggregating those up to a national index. Is that right? Is that fair?

Ori Heffetz

Yes. In principle, we could measure how an individual person is doing on different dimensions and then use their own weights to develop a personal index. That works in theory, doesn't work in practice, because no national statistical agency can go and survey each individual. So at the more aggregate level, we pool the responses of a lot of different people to create an index that is based on a group of people. If you survey different parts of the population or different demographic groups or different age groups, etc., you could get the voices of different groups heard or incorporated into those weights that you use to create the index.

Heather Howard 29:51

So you think with enough survey response, you could account for a geographic distribution. In a country like the U.S., for example, where there are different values clearly -- as we've seen in recent elections -- you think that your methods would allow you to account for that kind of geographic variation?

Ori Heffetz 30:11

In principle, you could. And I'll say, in a second, what I mean by principle. But in principle, just like we have political polls before the elections in every state, or sometimes in every county, you could do the same with our surveys, and you'll get the metrics and the weights for every region where you conduct the survey. I keep saying in principle, because in this sense, it's also a little bit similar to the political

polls. Those polls sometimes get it wrong, because not everybody wants to respond to the survey, and not everybody tells you the truth, and not everybody interprets the questions in the same way, and not everybody uses the response scale in the same way. So once you get into the mechanics of producing this data, there is a long list of caveats and a long list of challenges that we have been trying to tackle and to study through our research -- one by one, one challenge at a time. I don't want to give the impression that this is an easy thing, and tomorrow morning we could produce a perfect index. But this is something that we think deserves resources, and thinking, and work, and funding, to keep working on this research agenda and to keep improving things. As I said, in principle, this can be done. Now let's make it a reality.

Heather Howard 31:47

So let's go there. It's so great to have a researcher and a policymaker here, and I'd love to explore how your work is in conversation with each other. You're saying, Ori, that you're seeing this clearly as a long process -- that you've made significant progress, but there's more to come. Yanchun, do you find that your work is in conversation with Ori and his colleagues?

Yanchun Zhang

Of course. I think what Ori and his co-authors are doing is a very interesting piece of work, which could potentially be very important in providing information on the weights to be used. But on the other hand, almost any weighting procedure, including the one Ori proposed, or the current one that we are using in our office, attracts criticism. Because in a sense, it is arbitrary. Even, based on surveys, you ask questions and then the respondents can answer, but whether the true preference or choice is revealed by them is a different question. If you ask them, they would definitely place a lot of emphasis on the value of good education, good public health. But whether they are willing to pay for it, if they actually will pay for it, is a different question.

Then back to our HDI, we use a very simple weighting scheme, which is not perfect as I've stated a couple of times. We use one-third for each of the dimensions. It's a normative choice. It shows our value judgment. We think that health, education, and standard of living are equally important. It may not be consistent weight with every country's and every region's preferences, but all in all, we believe this is a good weighting system for our construction of HDI.

Heather Howard

Do you want to react to that, Ori?

Ori Heffetz 34:10

Yes, I agree with almost everything that Yanchun said. Every weighting scheme has weaknesses. In every weighting scheme there is moral judgment, an ethical judgment, a normative judgment. I think that starting off with 1/3, 1/3, 1/3, is not an unreasonable place to start. But it's also important to look at other ways to assess the trade-offs that people are making. Maybe in some places, and in some times, people don't view education, health, and income as equally important. You could imagine, for example, that in places where there is more income, but worse health, in fact, people would like to put more weight on health and less on income. And so it would be nice to get that voice of the people that I keep talking about -- to go and poll the people. Again, there is a long list of weaknesses with surveys.

Yanchun mentioned some, and I mentioned some, so surveys are not a silver bullet. But it would be good to collect these weights and to, at the very least, compare them with 1/3, 1/3, 1/3.

Also, it would be good to do something that your insurance office does all the time -- to conduct sensitivity analysis, to ask ourselves what the index would look like, or what the country rankings would look like, if we used a different weighting scheme. For example, if we used the weights of some region of the country, or some chunk of the population, or some other country. Nowadays we have the technical means and the wealth of data, so let's also collect this other data about the weights and about these other dimensions that are missing. And then ask ourselves if things look differently when you do these other things.

Yanchun Zhang 36:43

Absolutely. I think in terms of a research project, it's definitely worthy to pursue. But going back to the index construction... I always try to explain to our readers that HDI is not a purely statistical product. It's not a purely intellectual exercise. It is also a policy tool. So that's why the stability of the index is also important, and the methodology behind it shouldn't be changed frequently. But on the other hand, we are very open-minded about ongoing research that we could reference to make our methodology better. We often carry out these endeavors as separate research projects. We have developed, over the past 30 years, many complementary composite indices. When they become mature and well established, then we'll include them to the official offering of the Human Development Indices. So it's not only HDI, we have inequality-adjusted HDI, two gender indices -- gender development index and gender inequality index -- and also planetary pressures-adjusted HDI. So these are our efforts. For weighting schemes, we're definitely open to more research.

Ori Heffetz 38:11

I think that's a great example of how we would like our policymaking institutions to function when they are not under fire -- to be open to things that are more preliminary, that are kind of a crazy idea, that may not work. But they can do that research in the background and collaborate with academics, like us. Then only after these crazy ideas survive the test of time and prove themselves, etc, only then do we want to actually incorporate them into official statistics and base policy on them.

Heather Howard 38:56

So if we don't see next year that there's a whole new component, it doesn't mean that work is not going on behind the scenes. Work is always going on. Before we wrap, what's next, in terms of your research? Ori, is there anything you're excited about in this area that you're going to be focusing on in the next couple years?

Ori Heffetz 39:19

Yes. So for us, it's a two-pronged approach. On the one hand, we keep doing the research, to try to overcome all these challenges that I said before. Once you want to base policy on surveys, there's a lot of work to do to make the surveys more reliable, with better coverage and less bias, etc. So that's one thing. And on the other hand, we want to use the stuff we already understand -- to call out to policymakers, to statistical agencies, national statistical agencies, and policymakers all around the world. We need to tell them two things. First, let's start collecting more data on more dimensions of

wellbeing that traditionally we haven't collected data on. And secondly, let's start collecting those weights; let's have people make these hypothetical choices so that we will know their trade-offs, so that once we have the data, we will also have another way to assess the weights.

Heather Howard 40:26

Yanchun, any final words? Anything you're excited about in the next couple of years that you're working on?

Yanchun Zhang

For us, in terms of HDI, we really want to explore the following two questions. First, how do we go beyond basic capabilities? The current index only captures the basics. So the quality dimension definitely needs to enter the equation. The second thing we're working on is how to capture the agency. Agency is very important. So it's not only about your achievements, but also the freedom of choosing what life you want to lead, and then why you value such a life. So we try to have some refined index, or a new index that can help capture this important agency component of the original capabilities approach.

Heather Howard

Well, I'm really excited to hear about the evolution of measurement and the very exciting research you have underway. And maybe in a couple of years, we can have you back to talk about the new developments. This is an exciting space to be watching, and I so appreciate you joining us on the principles. Thank you.

Yanchun Zhang

Thank you so much, Heather.

Ori Heffetz

Thank you. Great to be here.