

15

WHAT THE GALLUP WORLD POLL COULD DO TO DEEPEN OUR UNDERSTANDING OF HAPPINESS IN DIFFERENT CULTURES

JEANNE L. TSAI

Do findings from the Gallup World Poll (GWP) tell us about people's happiness, well-being, and life satisfaction across the world, and if so, should they guide government policymaking? For the authors of the *World Happiness Report* (WHR) and at least some of its nine million readers, the answer is yes, and it is easy to see why.¹ First, these data capture people's feelings or subjective experiences, which have long been overlooked by economists and policymakers in their assessments of how nations are doing. Second, the GWP allows researchers to examine how people's feelings relate to environmental factors like GDP per capita, revealing that objective indicators like wealth do not tell the whole story about how people are feeling. Third, because the collaboration between the WHR and GWP has been going on for ten years, these data can reveal stability and change in happiness and well-being over time, particularly in response to life-changing events like the COVID-19 pandemic. Last, but certainly not least, the GWP samples 1,000 respondents from 150 nations to examine how people are doing in different parts of the world. As such, their data could reveal

important cultural similarities and differences in people's values and experiences of happiness.

Given the strengths of the GWP, why are the authors of *Against Happiness* so concerned? I will not repeat their many excellent points here, but their main argument—based on evidence from anthropology, philosophy, religious studies, sociology, and cross-cultural psychology—is this: happiness and well-being as currently defined and measured by the GWP may matter to its creators but may not matter as much to people across the world. If this is true, GWP data are limited in what they can tell us about how people are doing across the world, and using GWP data to shape government policies therefore would be premature and potentially harmful.

As a psychologist who has studied culture and emotion for thirty years, I agree with the authors' points. The good news is that there are ways to address the weaknesses of the GWP, which underlie many of the authors' concerns. Here I describe six ways to improve the GWP for studying happiness across cultures, and I describe how we have addressed these issues in our own work. The first three concern the measurement of happiness and well-being, while the last three concern the measurement of culture.

MEASURE VALUES AND IDEALS

By asking people to rate how close they feel to their best lives and to indicate whether they smiled, laughed, did something interesting, or felt worried or sad or angry during the previous day, the GWP focuses on respondents' actual experiences (their "actuals"). In doing so, the GWP assumes that the people in the 150 nations have similar "ideals." More specifically, the GWP

assumes that across the world, people want to lead their “best” lives, want to smile, laugh, and do interesting things and avoid feeling worried, sad, or angry. But what if many people in the study—and in the world—don’t hold these ideals?²

Fortunately, this question can be easily answered by also asking people how they ideally want to feel. By measuring respondents’ ideals, researchers can examine whether people in different regions of the world want to feel the same way about their lives, and if not, how they differ. Measuring people’s ideals can also help GWP researchers interpret the meaning of people’s actuals. Knowing whether people learned or did something interesting during the previous day matters a lot if people value learning or doing something interesting, and it matters very little if they do not. In other words, researchers can calibrate people’s actuals to their ideals. Finally, by measuring ideals and actuals, researchers can statistically account for social desirability biases. People are more likely to report feeling interest when they think it is desirable (or ideal) to do so and are less likely to report feeling anger when they think it is undesirable to do so. Thus, researchers can achieve a more accurate measure of people’s actuals by taking out the degree to which they overlap with their ideals. Last but not least, because ideals are powerful influences on people’s behavior in their own right, by measuring people’s ideals, the GWP would have a more complete view of what people care about.³

We know this because my colleagues, students, and I have been measuring people’s “actual affect” (the affective states they actually feel) and their “ideal affect” (the affective states they value and ideally want to feel) since 2000. Based on this work, we find that across the cultures studied, most people do indeed want to feel more positive than negative, and they want to feel more positive and less negative than they actually feel. Moreover,

among the positive states, people want to feel happier, more content, and more satisfied than other positive states. But against this backdrop of cultural similarities, we have also observed consistent cultural differences in the degree to which people value or ideally want to feel high arousal positive states like excitement, enthusiasm, and elation. European Americans valued these high-arousal positive states more than various East Asian groups, including Hong Kong Chinese, Taiwanese, Japanese, and South Koreans.⁴ Early on, we also found cultural differences in the degree to which people value or ideally want to feel low-arousal positive states like calm, relaxation, and serenity, with East Asians valuing these states more than European Americans.⁵ Most recently, these cultural differences in the valuation of these states have been less pronounced, perhaps because of recent world events.⁶ Cultures also differ in the degree to which they want to maximize positive and minimize negative states, with European Americans wanting to maximize the positive and minimize the negative more than Beijing and Hong Kong Chinese samples.⁷ These cultural differences are related to specific aspects of individualism and collectivism, respectively.⁸

Moreover, because in our studies we assessed actual and ideal affect in similar ways, we were able to examine the relative impact of each on a variety of behaviors. Our studies reveal that differences in ideal affect—above and beyond actual affect—are related to a whole host of behaviors, ranging from what consumer products people choose to what exercise they participate in, what physicians they prefer, what they post on social media, whom they view as friendly and whom they befriend, whom they hire and choose to lead, and even whom they share resources with.⁹

There are other ways in which assessing ideals can matter for well-being. Asking people to reflect on their ideals may remind

people what matters to them and ultimately help them to achieve those ideals. Learning about different ideals can broaden people's own views of a happy and meaningful life. Finally, recognizing that people differ in their ideals can be critical to understanding them better and minimizing unintended biases that may stem from different ideals.

INCLUDE FEELINGS THAT MATTER MORE IN COLLECTIVISTIC CULTURES AND OTHER RELIGIOUS TRADITIONS

People create products that reflect their cultural ideals. For instance, U.S. children storybooks, women's magazines, and even leaders' official website photos contain more open toothy "excited" smiles and fewer closed "calm" smiles than East Asian storybooks, magazines, and leaders' official website photos.¹⁰ The 2022 WHR is no different: photos in the report contained three times more excited than calm smiles!

Similarly, we have observed that Western well-being inventories contain more high-arousal positive than low-arousal positive content, presumably because they were developed by Western clinicians and scientists who value high arousal positive states more.¹¹ But what about the content of the GWP? As mentioned earlier, to assess positive emotions, the GWP typically asks respondents if they smiled, laughed, or did something interesting the day before, which assumes that when people feel good, they feel these emotions and express them in these specific ways. There are cultural differences in the emotions that people associate with well-being as well as cultural differences in whether people savor and express their positive emotions by smiling and laughing (vs. dampen and suppress their positive

emotions).¹² Therefore, it is possible that the GWP does not adequately capture the well-being of individuals from cultures that have different ideals or views of emotion, especially more collectivistic ones.

In response to this concern, the most recent GWP broadened its measures of positive emotion to include measures of balance, peace, and calm states not only valued in many East Asian contexts but many Buddhist traditions more specifically.¹³ This is a step in the right direction. Unfortunately, the GWP primarily focused on the degree to which respondents actually felt these states and included only one measure of ideals (“Would you prefer an exciting or calm life?”). As argued before, in order to evaluate people’s reports of actual balance and peace, we need to know how much they value those states. Moreover, in order to be comparable, actuals and ideals have to be measured in the same way. Despite this, consistent with our early findings, the authors still observed that more respondents in East Asian countries preferred a calm vs. exciting life compared to respondents in the United States and Canada.¹⁴ Moreover, they observed that during a world pandemic, most people across the world preferred a calm over an exciting life. Whether these findings will generalize to nonpandemic times remains to be seen, but the findings illustrate another reason to include other feelings that matter to other parts of the world and other traditions: we may discover that other states matter to us more than we realize.

Similar arguments can be made for negative emotion. The GWP focuses on three negative states—worry, sadness, and anger—and assumes that people want to avoid feeling these specific states. Once again, we know from our work that people and cultures vary in the degree to which people want to avoid feeling negative overall as well as the degree to which they want to feel and avoid specific negative emotions.¹⁵ European Americans,

for example, want to avoid feeling negative more than Germans, and this is linked to their expressions of sympathy.¹⁶ Taiwanese participants value shame, guilt, and fear more than Canadians do.¹⁷ Moreover, views of negative emotion have implications for the links between the actual experience of these negative states and health.¹⁸

Future GWP might include other states that have been associated with collectivism, given the prevalence of collectivistic values across the world. For instance, Shinobu Kitayama, Batja Mesquita, and Mayumi Karasawa have demonstrated that people in cultures that are more collectivistic like Japan experience more socially engaging positive (e.g., feelings of friendliness and respect) and negative (e.g., feelings of pity and shame) emotions; therefore, these would be good states to include in future polls.¹⁹

VALIDATE SELF-REPORTS WITH REAL-WORLD BEHAVIORS

The WHR primarily relies on self-report data from the GWP to assess how people are doing across the world. Although self-report is one of the best ways of assessing people's subjective experiences, there are a host of problems with it, including social desirability biases, response style biases, and lack of awareness or insight. One of the most important concerns is whether people's self-reports of how they are doing reflect anything about people's everyday lives. For instance, are people's assessments of their lives truly a reflection of what their lives are like? Does a measure of whether people smiled, laughed, or did something interesting the day before say anything about people's lives overall? To answer these questions, it is critical to validate these

self-reports with other behavioral and biological measures of how people are doing.

The 2022 WHR has already moved in this direction by examining whether the affective content of Twitter posts during the COVID-19 pandemic reflects the same patterns as the GWP data. The question is what the affective content in the posts reflects. In our own work comparing the affective content of U.S. and Japanese Twitter posts, we observe patterns that suggest that people are posting content that reflects their ideal affect more than their actual affect.²⁰ By measuring ideals and actuals, GWP might be able to determine whether this is true for its respondents as well.

As described earlier, we have employed a variety of behavioral and biological measures to examine the behavioral expressions of cultural differences in ideal affect, as well as the neural mechanisms that support these behaviors. For instance, in one study, based on their self-reports of ideal affect, European Americans valued excited (high-arousal positive) states vs. calm (low-arousal positive) states less than Chinese. Within cultures, European Americans valued excited and calm states similarly, whereas Chinese value calm more than excited states. To examine how these cultural differences influence people's immediate responses to excited vs. calm facial expressions, we combined a facial rating task with functional magnetic resonance imaging.²¹ During scanning, European American and Chinese females viewed and rated excited and calm faces that varied by race (white, Asian) and sex (male, female). As predicted, European Americans showed greater activity in brain circuits associated with affect and reward (bilateral ventral striatum, left caudate) while viewing excited vs calm faces than did Chinese. Within cultures, European Americans responded to excited vs. calm faces similarly, whereas Chinese showed greater activity in these circuits

in response to calm vs. excited expressions regardless of targets' race or sex. Across cultures, greater ventral striatal activity while viewing excited vs. calm faces predicted greater preference for excited vs. calm faces several months later. These findings not only provide neural evidence that people find viewing the specific positive facial expressions that match their culture's ideal affect to be rewarding and relevant but also validate the cultural differences that have emerged in self-reports.

DEMONSTRATE MEASUREMENT EQUIVALENCE

As the authors of *Against Happiness* argue, the questions included in the WHR may have different meanings in different cultures, especially if the translations are not accurate. As a result, comparing ratings of these happiness metrics across cultures may ultimately be meaningless.

Indeed, this is a fundamental issue that cross-cultural psychologists have confronted for decades, which is why standard practice is to demonstrate measurement equivalence *before* comparing group responses.²² Demonstrating measurement equivalence requires showing that responses to questions on a survey relate to each other in similar ways across the cultures sampled. For instance, in the GWP, are smiling, laughing, and doing something interesting related to each other for each of the nations sampled? They should be if they are truly assessing positive emotion in each culture. If so, then comparisons of positive emotion across nations are warranted. But if they aren't, then they should be compared separately across nations and not referred to as "positive emotion." Similarly, do "balance" and "harmony" have the same associations in the 150 nations sampled?

Because we have established measurement equivalence for our measures of actual and ideal affect for the cultural samples we have studied, we are confident that we can compare actual and ideal affect between these cultures. Thus, we know that for our participants, how people want to feel differs from how they actually feel, and there are cultural differences in how people ideally want to feel.

ACCOUNT FOR CULTURAL DIFFERENCES IN RESPONSE STYLES

A central aspect of the GWP and WHR is the ranking of nations by average level of life satisfaction on a 10-point scale. Rating scales like this one allow nuance and variability more than simple yes or no questions, but they have issues, too. For instance, decades of research in cross-cultural psychology have revealed cultural differences in how people respond to rating scales.²³ While East Asians tend to use the middle of the scales, Americans tend to use the extreme ends of the scales. This is not a problem when comparing ratings within cultures, but across cultures it means that one cannot know whether the observed national differences in life satisfaction reflect real differences in life satisfaction or more general differences in how much people use the middle vs. top and bottom ends of the scale. Yes and no questions, which the GWP uses to assess positive and negative emotions, are similarly sensitive to cultural differences in “acquiescence bias,” or their willingness to provide affirmative answers.²⁴

In our own work, we ask participants to use a 5-point rating scale to indicate how much they actually feel thirty-eight different states that vary in terms of valence and arousal on average

and then how much they ideally want to feel the same thirty-eight states on average. Although we sometimes reduce the number of items depending on our purposes, these thirty-eight states include filler items or items that are not central to our hypotheses so that we can assess and then account for differences in response styles. One popular way of doing this is "ipsatizing," or calculating standardized scores for each individual. We then compare the ipsatized scores across cultural groups. Although this method has its limitations, it addresses response style biases and allows us to infer more confidently that the observed cultural differences reflect specific differences in ideal affect and not more general differences in response styles.²⁵

PROVIDE MORE CULTURAL DEPTH (EVEN AT THE COST OF NATIONAL BREADTH)

The GWP samples 1,000 respondents from 150 nations ranging from Finland to Afghanistan. This is not an easy task, especially assuming that the GWP employs translators and researchers in each of these nations to collect these data. If you have the resources, why wouldn't you collect data from as many nations as possible?

From a scientific perspective, one reason is that this approach almost always comes at the cost of understanding of any particular nation or collection of nations well. For instance, at least in the most recent WHR, almost no information is provided about the specific cultural ideas and practices represented in this collection of 150 nations that might be related to different conceptions of happiness or well-being. This is not specific to the GWP; it is almost always the case that in studies that include

more than a handful of nations, the more nations the researchers include in their studies, the less cultural the study becomes. Although empirical studies are by definition reductive, with respect to culture, the GWP takes this to an extreme.

This is exactly why when I began studying culture influences on emotion, I focused on two ethnic groups: European Americans and Chinese Americans.²⁶ I chose these groups because they were differentially influenced by and oriented to “American” and “Chinese” cultures, cultures that differed in individualism and collectivism, which scholars had predicted would shape a variety of psychological processes, including emotion.²⁷ By focusing on these two groups, my team and I could also consider and quantify the variation within each cultural group in terms not only of individuals’ cultural orientation but also in their endorsement of individualistic and collectivistic values. We also used specific recruitment criteria to ensure that we were recruiting individuals who were oriented to the cultures of interest (e.g., typically first- and second-generation Chinese Americans and third- or higher generation European Americans).

In subsequent studies for various theoretically and empirical driven reasons, we broadened our samples to include other East Asian Americans, then participants living in Taiwan and Hong Kong, then mainland China, and then other East Asian cultures (Japan, Korea) as well as other Western cultures (England, France, Germany) and other collectivistic cultures (Mexico).²⁸ By choosing nations for cultural reasons, we believe that we have been able to increase our national breadth without completely compromising cultural depth (although some anthropologists might disagree).

This focused approach has also allowed us to study cultural differences in ideal affect more deeply. Specifically, we have been able to test our predictions about the sources of cultural

differences in ideal affect, examine how cultural differences in ideal affect are reflected in popular media, look into age differences in ideal affect, explore the interpersonal consequences of cultural differences in ideal affect, reveal the neural mechanisms underlying cultural differences in ideal affect, study the effects of differences in ideal affect in employment and health settings, and most recently to chart changes in ideal affect over time.²⁹ We would not have been able to do any of this if we did not have a solid understanding of what ideal affect looked like in the specific cultural samples we have studied.

By starting out with U.S.–East Asian comparisons, we did not mean to imply that they were the only cultures that vary in the valuation of high and low arousal positive states, and I certainly believe that we need to know more about the ideal affect of other parts of the world, which other scholars have fortunately initiated.³⁰ In my own lab, we have started studying ideal affect in Turkey and India, with collaborators who know these cultures well, in part because these cultures (and the subcultures within them) have different expressions and combinations of individualism and collectivism that should show different patterns from those we have observed in our samples.

WHY IMPLEMENT THESE CHANGES?

Implementing these six changes would require additional time and resources, but they are necessary if the authors of the WHR really want to understand the meaning of happiness and well-being in different cultures. Armed with better data, the WHR would have more to say to government policymakers about how their constituents feel compared to how they ideally want to feel. As I read the WHR, I was heartened to learn that the WHR

and GWP have evolved over time, and that at least some of its authors share the points raised here and by the authors of *Against Happiness*. Perhaps by implementing the six changes described here, the GWP will begin to define and measure happiness in ways that help get us all closer to understanding happiness and well-being across the world.