Three questions about happiness

John F. Helliwell

Abstract
Frijters, Clark, Krekel and Layard (2019) makes a strong historical case for treating happiness as a primary measure of human welfare. They argue that it is now time to revamp the science of policy-making to re-establish this central tradition, building on recent progress in measuring and understanding subjective well-being. I agree with them. There are three key questions raised in or by their paper that need further evidence. I shall try to address these in turn. The questions relate to how to measure happiness, how to measure and deal with inequality, and how to take due account of the social context of well-being, including the need to achieve consistency between individual and societal happiness.

Jel codes: I31, D63, D64

* Accepted version of a paper scheduled to appear, along with the paper by Frijters et al (2019) and other commentaries in a special issue of Behavioural Public Policy. I am grateful for the assistance and advice of Lara Aknin, Max Norton, and Shun Wang.
Frijters, Clark, Krekel and Layard (2019) argue that it is now time to revamp the science of policy-making, building on recent progress in measuring and understanding subjective well-being. There are three key questions raised in or by their paper that need further evidence: how to measure happiness, how to measure and deal with inequality, and how to take due account of the social context of well-being, including the need to achieve consistency between individual and societal happiness.

How best to measure happiness to support public policies?

Frijters et al. (2019) recognize but do not adjudicate among various measures of subjective well-being and indices constructed from a range of indicator variables. In my view, there is now enough evidence in hand to establish clear relationships among the available measures. The central role is and should be played, as argued by Aristotle, by people’s own evaluations of the quality of their lives. When we were explaining in the World Happiness Report 2015 why we had no hesitation in using people’s own life evaluations instead of some constructed index of social and economic indicators, here are the four main reasons we offered:

“First, we attach fundamental importance to the evaluations that people make of their own lives. This gives them a reality and power that no expert-constructed index could ever have. For a report that strives for objectivity, it is very important that the rankings depend entirely on the basic data collected from population-based samples of individuals, and not at all on what we think might or should influence the quality of their lives. Thus the average scores simply reflect what individual respondents report to the Gallup World Poll surveyors. The Report editors have no power to influence the averages beyond the choice of the number of survey years to use to establish sufficiently large samples.

Second, the fact that life evaluations represent primary new knowledge about the value people attach to their lives means we can use the data as a basis for research designed to show what helps to support better lives.

Third, the fact that our data come from population-based samples in each country means that we can calculate and present confidence regions about our estimates, thus providing a way to see if the rankings are based on differences big enough, or not, to be statistically meaningful. If a number of adjacent ranked countries all have values well within the sampling range of variance, then it can be concluded that they deserve to be treated as having statistically equivalent average life evaluations.

Fourth, all of the alternative indexes depend importantly, but to an unknown extent, on the index-makers’ opinions about what is important. This uncertainty makes it hard to treat such an index as an overall measure of well-being or even to work out the extent to which variations in individual components are affecting overall scores. Even where this decomposition is done, there is no way of establishing its validity, since the index itself is just the sum of its parts, and not an independent measure of well-being.” (Helliwell,
Huang & Wang 2015, 19-20)

Should positive and negative emotions be seen as competing with life evaluations as measures of human happiness, and possibly be preferred to them, as argued at different times by both Bentham and Kahneman (Kahneman et al 1997)? Or are positive emotions instead better seen, as argued by Aristotle, as contributing to life evaluations rather than being sufficient measures on their own? The empirical evidence from the Gallup World Poll, as shown for example in Table 2.1 of the World Happiness Report 2019, is that positive emotions (but not the absence of negative ones) contribute to life evaluations above and beyond the six main variables (GDP per capita, healthy life expectancy, social support, freedom, generosity and trust) used to represent different aspects of life. This finding was predicted by Aristotle and demonstrated more recently by Fredrickson (2004). Introducing positive emotions into the explanation of life evaluations significantly lowers the coefficients on a sense of freedom and having someone to count on in times of trouble, suggesting that positive emotions are likely mediators for how those variables influence life evaluations (e.g. Lyubomirsky & Layous 2013). Analogously, positive emotions are strongly correlated with both of those variables, while negative emotions are most strongly correlated with absence of social support and presence of corruption.

What about measures of a purposeful life, sometimes measured and treated as primary indicators of well-being? For example, a life purpose question is included as one of four well-being measures regularly collected by the UK Office for National Statistics (along with life satisfaction, and two emotional measures- happiness yesterday and anxiety yesterday) as the central statistical base for the measurement of national well-being. The resulting measures are much more highly correlated with life evaluations than they are with either positive or negative emotions. Interestingly, having children at home, often found to be negatively correlated with life evaluations, (presumably as contributors to excess demands on bandwidth sometimes scarce at those ages when family and career pressures combine) is positively correlated with the answers to the life purpose question (for experimental evidence to this effect, see Nelson et al. 2013). In a similar vein, it might be expected, when further analysis is possible, that life purpose contributes to positive overall life evaluations, as Aristotle saw to be likely, and just as already found for a sense of freedom to make key life decisions. A sense of purpose is thus best seen, as is a sense of freedom, not as an independent overall measure of well-being but rather as a key contributing factor, a right-hand-side variable likely to be useful in explaining overall life evaluations.

Evidence continues to accumulate showing that life evaluations, positive affect and negative affect, long thought to capture different aspects of well-being (Diener et al 1999) are now seen to do so in a consistent manner. Of these three ways of measuring subjective well-being, life evaluations provide the most encompassing indicator of the quality of life. The three most common forms of life evaluation are life satisfaction questions of the sort advocated by the OECD (2013), questions asking how happy people are with their lives as a whole, and the Cantril ladder question used in the Gallup World Poll, wherein people are asked to think of their lives as a ladder, with the worst possible
life for them as a 0 and the best as a 10. Where these different life evaluation questions have been asked of the same respondents in the same survey (SWL and the Cantril ladder in one year of the Gallup World Poll, and SWL and happy with life regularly in the European Social Surveys) they tell structurally equivalent stories about the correlates of a good life, even though the means and shapes of the distributions of the answers differ among the questions.

Although positive and negative emotions are not suitable as overall measures of the quality of life, they are important in other key ways. Their short-term nature makes them natural to use as outcome variables in experimental and time-use studies (Kahneman et al. 2004). They are also the most important means of measuring weekend effects, which show up in daily emotional reports but not in life evaluations (Helliwell & Wang 2015, Figure 2), thus demonstrating the congruence of the two types of measure. Furthermore, the frequency of a number of positive and negative emotions (especially positive emotions, e.g. Cohen & Pressman 2006) has been found to predict a number of future health outcomes, including mortality from both sickness and suicide. Only recently is it becoming more common to collect positive emotions and life evaluations as part of the data used to track and diagnose physical health, so it is still too early to tell the relative importance of emotions and life evaluations as predictors of various health outcomes and, of course, vice versa.

Some have argued that it is misleading to use ‘happiness’ as a generic term to cover subjective well-being more generally. While ‘subjective well-being’ is more precise, it simply does not have the convening power of ‘happiness’. The main linguistic argument for using happiness in a broader generic role is that happiness plays two important roles within the science of well-being, appearing once as a prototypical positive emotion and again as part of a cognitive life evaluation question. This double use has sometimes been used to argue that there is no coherent structure to happiness responses. The converse argument made in the World Happiness Reports (also Helliwell & Aknin 2018) is that this double usage helps to justify using happiness in a generic role, as long as the alternative meanings are clearly understood and credibly related. Evidence from a growing number of large scale surveys shows that the answers to questions asking about the emotion of happiness differ from answers to judgmental questions asking about a person’s happiness with life as a whole in exactly the ways that theory would suggest. Answers to questions about the emotion of happiness relate well to what is happening at the moment. Evaluative answers, in response to questions about life as a whole, are supported by positive emotions, as noted above, but also driven much more, than are answers to questions about emotions, by a variety of life circumstances, including income, health and social trust.

Only life evaluations, among all the individual or composite indicators listed by Frijters et al. (2019) meet two primary tests for an encompassing measure. First, they are themselves overall assessments of the quality of life, without needing any further construction or manipulation. Second, since they are primary measures and all-encompassing in their scope, they provide the research base for answering the fundamental questions about the quality of life – what does tend to lead to a better life, as
seen by those doing the living?

How should inequality be seen through the happiness lens?

How should well-being policies recognize and respond to the large and in many places growing gaps between the haves and have-nots? Frijters et al. (2019) recognize the importance of distribution, but reach no specific conclusions. It is possible to go further. If life evaluations are the appropriate measure of individual and average well-being, then the spread of this distribution should be the primary measure of inequality. And if people on average share the preferences of those who have advocated a range of inequality-averse social welfare functions, we might expect to find that people are happier living in communities or nations where differences in well-being between the haves and have-nots are smaller. Furthermore, if that should turn out to be the case, then the weight that typical citizens attach to well-being inequality can be used to evaluate policies that affect the distribution of well-being. What does the evidence show?

Several international and domestic well-being surveys have been used to show that well-being inequality, as measured by the standard deviation of the distribution of Cantril ladder scores, had a larger and more pervasive negative link to national average scores than did measures of income inequality (Goff et al. 2018). This effect was seen to persist even using several ways to offset the negative bias imparted by top-end truncation of the ladder scale, and to be stronger for those individuals who reported a wish to see inequality reduced. Subsequent research has found that ratios of the percentiles of the national distributions of responses to the Cantril ladder question, which are better insulated against possible end-point effects, have an even stronger negative partial correlation with national average happiness levels. Average life evaluations are therefore already adjusted for the welfare effects of inequality, according to how the average respondent’s happiness is correlated with the percentile ratio. There is thus no need to further adjust the mean of the well-being distribution to reflect inequality, as this is already done by survey respondents. This analysis illustrates another advantage of having a direct measure of well-being. Those who have used GDP per capita as a proxy measure of welfare have had to adjust for inequality by assuming some form for the underlying welfare function to derive an implicit trade-off between average income and the inequality of its distribution (e.g. Atkinson & Brandolini 2010). By contrast, having a directly measured umbrella indicator for well-being permits direct estimation to replace assumptions about the extent to which inequality, whether of income or well-being itself, affects welfare.

Policies that affect happiness will therefore have total effects on well-being that reflect their combined influence on the levels and distribution of well-being. In this way, policies that have their primary impact on people who are most disadvantaged will increase average well-being by more than just their effects flowing through the well-being of the disadvantaged themselves. This is because others are happier living in these new circumstances with more equally distributed well-being. Beyond this, there is also evidence that policies having a positive effect on overall social trust provide larger well-
being gains for individuals who are unemployed, in ill-health, or subject to discrimination (Helliwell, Huang & Wang 2018), all of which are conditions more prevalent among those at the bottom end of the well-being distribution. Individuals within a society and among societies differ in the extent to which their happiness is affected by inequality (Alesina et al. 2014). Goff et al. (2018) have found that those whose life evaluations are most negatively affected by happiness inequality are also in favour of policies directed to reduce inequalities. Average life evaluations are already adjusted for the welfare effects of inequality since they account for individuals’ aversion to inequality, estimated to be of substantial size (Goff et al. 2018). These estimates in turn can be used to help rank policies that have differing impacts on the inequality of well-being.

A clear focus on well-being inequality is by this means likely to deliver a more efficient set of policy options to be considered, and a better way of ranking their overall effects on the level and distribution of happiness.

**Accounting for the social context of well-being**

Another advantage of having a direct measure of utility is that it helps establish the relative importance of various factors influencing well-being, and perhaps thereby to increase attention paid to factors not previously taken into account. Such is the case for variables measuring the nature and quality of the social context. Early happiness research did establish the importance of strong and satisfactory relationships with family and friends (Diener & Seligman 2002). However, there were few systematic attempts to explore the comparable social contexts at work, in the community, and more generally in the ways in which people interact with each other and with their governments. Early surveys tended to be relatively weak in relevant social measures, with some national and international social surveys providing notable exceptions.

When surveys did permit the relevant social contexts to be assessed, they were quickly seen to be of first-order importance, thereby triggering interest in more and better data. One key set of social variables, driven by interest in social capital (Putnam 1994, 1995, OECD 2001), relates to various aspects of trust, and in individual and community-wide measures of social engagement. Where separate questions have been asked about the climate of trust in the workplace, the neighbourhood, the police, and in society in general, they all appear to be independently important (Helliwell & Putnam 2004). And two-level models show that social trust and sociability have positive contextual effects above and beyond their individual-level effects (Bjørnskov 2008). Positive contextual effects for measures of sociability tend to appear if the local community is defined on a sufficiently fine-grained basis (Han et al. 2013).

**Valuing social capital**

Using life satisfaction equations to estimate an equivalent financial value (via compensating differentials) implies high values for a trustworthy environment. At the global level, these have been estimated to average 20% of World Bank estimates of total wealth (comprising stocks of produced, natural, human and social capital for 132
countries with available data for social trust and capital stocks; Hamilton et al, 2017).

In the workplace, to have a job where trust in management is one point higher on a ten-point scale has been estimated to be worth the equivalent, in terms of life satisfaction, to a 30% higher income (Helliwell & Huang 2011).

One especially interesting trust question asks respondents how likely their lost wallet would be to be returned if found, alternatively, by a stranger, a neighbour or a police officer. To return a lost wallet is a generous prosocial act, going far beyond any formal definition of trustworthiness. As such, it is no surprise that people are much happier living in communities or countries where wallet return is thought to be likely. Wallet return questions were asked in 132 countries in the 2006 Gallup World Poll. Those with a high expected wallet return were significantly happier, by an amount more than equivalent to a doubling of income (Helliwell & Wang 2011, Table 2-d). These expectations of wallet return reflect underlying realities, as the average national rates of expected wallet return, if found by a stranger, are highly correlated (r=0.83, p<.001), for the 16 countries in both samples, with the actual return of wallets in a recent large experimental study (Cohn et al. 2019).

**Valuing the community**

Having large scale surveys that include some measures of social connection opens doors to measuring the importance of having a sense of belonging in one’s local community. This too has turned out to be of first-order importance, both at the individual level and in explaining differences among communities in their average life evaluations. At the individual level, having a strong sense of belonging to one’s local community is strongly supportive of life satisfaction even after taking into account strong positive correlations with several types of trust (Helliwell & Wang 2011, Table 3). Estimates from the Canadian General Social Survey show that differences among respondents in their sense of belonging to their local community explains more of their life evaluations than does household income. The sense of community belonging is itself estimated to depend on several measures of community strength, including several types of trust, the extent and warmth of social connections, and time spent living in the neighbourhood (Helliwell & Barrington-Leigh 2012).

Across communities, differences in average sense of community belonging are even more important correlates of average life satisfaction. They explain (in terms of partial correlations measured by standardized betas) four times as much as do inter-community differences in income for life satisfaction among 1215 equally sized neighbourhoods and communities defined by natural and built community boundaries (Helliwell, Shiplett & Barrington-Leigh 2019).

**Valuing and enabling prosocial behavior**

There is also a closely related research stream, with strong implications for the design and delivery of public policies, that studies the happiness consequences of prosocial
motivations and behavior (Aknin et al 2019). Some evidence suggests that people who volunteer are happier and healthier than non-volunteers, based on cross-sectional surveys in many countries (Borgonovi 2008, Haski-Leventhal 2009). In the same vein, Brown and colleagues (2003) found that mortality rates were lower for individuals who provided instrumental support to family, friends, and neighbours as well as emotional support to their spouse (also Schwartz & Sendor 1999; Thomas, 2009). Engaging in other forms of kindness, such as financial generosity, appears to have similar benefits. Charitable activity is also correlated with higher life evaluations, as individuals who recently donated money to charity reported higher levels of life satisfaction, in most countries around the globe, than non-donors (Aknin, Barrington-Leigh et al., 2013).

Establishing causal directions linking prosocial behavior and happiness

Since causation in these cases is likely positive in both directions, experiments have been used to establish causal linkages from prosocial actions to subjective well-being, where the latter is usually measured by positive affect, given the typically small scale and temporary nature of the experimental interventions. Participants randomly assigned to engage in kind acts for other people or for the world, reported greater positive emotions and decreased negative emotions, compared to controls (Nelson et al. 2016). Similarly, individuals randomly assigned to provide emotional support to partners displayed greater activation in pleasure centers of the brain than individuals not able to support their partners (Inagaki & Eisenberg 2012). Moreover, people randomly assigned to spend money on others, called prosocial spending, report greater positive emotion than people randomly assigned to spend money on themselves (Dunn, Aknin & Norton 2008, 2014). This latter finding has been detected in hundreds of people from rich and poor countries around the globe (Aknin, Barrington-Leigh et al. 2013), as well as among ex-offenders (Hanniball, Aknin, Douglas & Viljoen, 2019). Even children under the age of two smile more when giving treats away to others than when receiving treats themselves (Aknin, Hamlin & Dunn 2012).

When do prosocial acts work best to produce happiness?

Experiments have also been used to study the conditions under which prosocial acts are more likely to increase happiness, and the results confirm that social capital and pro-social behavior are mutually supportive. For example, pro-social actions have been shown to have greater happiness benefits if they facilitate social interactions (Aknin, Dunn, Sandstrom & Norton, 2013), if they are undertaken for altruistic rather than selfish motivations (Konrath, Fuhrel-Forbis, Lou & Brown 2012; Veerasamy, Sambasivan & Kumar 2015), if they are done voluntarily (Weinstein & Ryan 2010), and if they are thought to be of greater benefit for others.

Happiness in its turn has long been known to incite pro-social actions (Isen & Levin 1972). Although this makes unpacking the causal chain more complicated, such
positive feedbacks create desirable multiplier effects from interventions that enable pro-social behavior to flourish.

Thus prosocial acts produce greater happiness benefits in conditions where they are supported by strong social capital, as measured by trust and positive social interactions. And the resulting happiness improves the climate for subsequent prosocial behavior. This positive synergy provides many opportunities to improve the design and delivery of public services.

**References**


the globe. Science, 365, 70-73.


