



Happily Distant or Bitter Medicine? The Impact of Social Distancing Preferences, Behavior, and Emotional Costs on Subjective Wellbeing During the Epidemic

Sarah Kelley¹ · M. D. R. Evans² · Jonathan Kelley³

Received: 1 December 2020 / Accepted: 19 February 2022 / Published online: 4 March 2023
© The International Society for Quality-of-Life Studies (ISQOLS) and Springer Nature B.V. 2023

Abstract

To inhibit the spread of COVID-19 Public health officials stress, and governments often require, restrictions on social interaction ("social distancing"). While the medical benefits are clear, important questions remain about these measures' downsides: How bitter is this medicine? Ten large non-probability internet-based surveys between April and November 2020, weighted statistically to reflect the US population in age, education, and religious background and excluding respondents who even occasionally role-played rather than giving their own true views; $N=6,223$. Pre-epidemic data from 2017–2019, $N=4,032$. Reliable multiple-item scales including subjective wellbeing (2 European Quality of Life Survey items, Cronbach's $\alpha=.85$); distancing attitudes (5 items, $\alpha=.87$); distancing behavior e.g., standing 6' apart in public (5 items, $\alpha=.80$); emotional cost of distancing and restrictions on social interaction (8–12 items, $\alpha=.94$); and an extensive suite of controls (19 variables). Descriptive statistics, OLS regression, structural equation models. Subjective wellbeing is greater for those who approve of distancing, for those who practice distancing, and particularly for those whose distancing attitudes and behavior are congruent, either both in favor or both opposed (multiplicative interaction). The emotional cost of distancing is strongly tied to wellbeing and is heterogeneous, with some disliking distancing much more than others. An SEM model suggests causality: that emotional costs strongly reduce wellbeing but not vice-versa. During the epidemic, COVID issues constitute two of the top 5 influences on wellbeing, behind only subjective health and religious belief and tied with income. All this is net of family background, religious origins, age, ethnicity, race, gender, rural residence, education, occupational status, marriage, unemployment, income, health, religion, and political party.

✉ Sarah Kelley
sarahkelley1759@gmail.com

Extended author information available on the last page of the article

Introduction

The coronavirus epidemic has wrought many changes in our social life and customs, but we are only beginning to understand their implications for wellbeing. Multiple previously unusual behaviors are suddenly mandated or recommended by public health authorities and governments that would have been thought bizarre or even dangerous just six months before the epidemic. For example, people who decline to shake hands would have been seen as neurotically fearful; people wearing masks into a store would likely have been seen as robbers; people who shun parties might have been seen as painfully shy; people who avoid in-person visits with friends and instead only visit electronically might have been seen as a recluse. Similarly, staying home and hardly ever going out might have signaled agoraphobia!¹ As late as mid-March 2020, a faculty member was scolded by a responsible and well-informed chair for wanting to hold class remotely!²

While these new public health measures are necessary to inhibit the spread of coronavirus, they nonetheless have little-studied impacts on subjective wellbeing. Much public discourse suggests that adopting these behaviors has little or no emotional cost and is merely a matter of habit. But that is a heretofore untested assumption, and, as our findings will show, that assumption is grossly incorrect.

Our purpose here is to investigate the norms that people hold about these behaviors; the prevalence of precautionary behaviors; the pleasantness or unpleasantness of performing the precautionary behaviors – the emotional valence of the behavior itself; and the connections of these behaviors with subjective wellbeing. This analysis is done using ordinary least squares regression. In addition, we assess the degree to which emotional feelings about precautionary behaviors ("emotional costs") influence subjective wellbeing and the degree to which influence flows in the opposite with a structural equation model allowing reciprocal causation between subjective wellbeing and emotional costs. We also examine the degree to which distancing behavior's connection with subjective wellbeing is contingent upon one's attitude about distancing, via augmenting the main OLS model with a multiplicative interaction between distancing behavior and distancing attitude.

Our epidemic era data are from a series of ten nationwide cross-sectional internet surveys conducted in the United States between April and November 2020. Our pre-epidemic data are from 2017 to early 2019.

To improve the precision of the estimates of statistical relationships and reduce the "noise" caused by random measurement error, we use multiple-item scales wherever possible. Subjective wellbeing is measured by the average of two standard items

¹ The term "social distancing" is often used loosely by the general public and social media to encompass both social deprivation behaviors (refraining from in-person contact with people to whom you are socially connected) and other behaviors sometimes termed "public" or "civic" or "casual", such standing 6 feet apart from others, refraining from shaking hands, and, on a volte-face part-way through the epidemic, wearing a mask. All of these seem to cohere in terms of behavior – there is one "distancing cultural practice" that they all tap, but the patterning of emotional costs is more complex, as discussed in the Methods section and Supplementary Material Appendix EMOTIONAL COSTS.

² Personal Communication A1, anonymized to protect privacy.

from the European Quality of Life Surveys (EQLS): happiness and life satisfaction. Our social distancing/ COVID-precautionary behavior scale (KEK-2 Distancing Behavior) is relatively new, so although its measurement properties are assessed in detail elsewhere (Evans et al., 2021), we briefly review them here. Here we introduce a measure of favorable (or unfavorable) attitude toward the precautionary behaviors (KEK-3 Attitude to Social Distancing) and two scales of the emotional cost of precautionary behavior, one focusing on refraining from social interaction (e.g., not visiting friends) and the other focusing on civic/ casual precautions (e.g., standing 6 feet away from other people in public). Comprehensive details on data, measurement, and models are in a series of Supplementary Material Appendices.³

Theory and Prior Research

The question of how COVID-19 precautionary behaviors/ distancing relate to subjective wellbeing is complex. In terms of motivation, distancing is both a set of personal health behaviors that directly protects oneself and one's household from contagion and a set of altruistic acts that reduces contagion to protect fellow citizens and possibly to 'flatten the curve'. From the 'personal health' standpoint, a widespread, possibly bidirectional, relationship has been seen between standard health promotion behaviors (e.g., non-smoking, exercising, eating healthfully) and wellbeing (Grant et al., 2009). This would suggest that engaging in social distancing behaviors would increase wellbeing slightly from a general sense of working towards their own good health. Early evidence from China suggests that, in fact, individual decisions to self-quarantine were associated with greater happiness (Lu et al., 2021).

From the 'altruism' standpoint, there is also a broadly documented relationship between acting altruistically and happiness (e.g., Post, 2014). Of course, the causality is complex (do happier people help others or does helping others boost happiness?), but a growing body of experimental and longitudinal research (and the discovery of plausible biological and psychosocial mechanisms) suggests that altruism does likely have a causal impact on wellbeing (Post, 2014). For instance, a US longitudinal survey found volunteering effects on subjective wellbeing in 2004, controlling for 1992 subjective wellbeing (Piliavin & Siegl, 2007), a finding that has now been replicated in the UK (Lawton et al., 2021). These results suggest, similarly, that the altruistic aspect of COVID-19 precautionary behavior also should increase subjective wellbeing. Taken together, these results suggest that distancing would likely *increase* subjective wellbeing modestly though making one feel that one is taking care of one's health as well as feeling that one is making a social contribution by protecting others from getting sick.

³ Appendices: CLASSICAL MEASUREMENT MODEL, COLINEARITY, CORRELATIONS, EMOTIONAL COSTS, HOUSEHOLD COMPOSITION, INTERACTION, MARRIAGE, MISSING DATA, RECIPROCAL EFFECTS, SAMPLE DESCRIPTION, SEVERITY OF COVID, VERTICAL OCCUPATIONAL STATUS, and WEIGHTING & ROLE-PLAYING BIAS. They are in Supplementary Materials (in alphabetical order).

On the other hand, a deep literature on happiness and wellbeing suggests we have reason to be concerned that distancing – which drastically reduces the amount and variety of social contact people have – will reduce wellbeing. Even pre-pandemic, social isolation and loneliness were increasingly recognized as growing public health concerns (e.g., Holt-Lunstad, 2017) and perceived isolation (e.g., Santini et al., 2020) is widely associated with worse mental health outcomes during the pandemic. This suggests that social deprivation distancing would likely substantially reduce wellbeing and that the public/civic/casual precautionary behaviors might, as well. In fact, a framework for ending lockdowns has been proposed that explicitly seeks to take into account the effect of the lockdowns on wellbeing. They try to calculate the wellbeing effects of the lockdown itself, loss of income and other variables and use this to calculate the loss of wellbeing-adjusted life-years under different policy scenarios (Greyling et al., 2021b; Layard et al., 2020).

An additional complication to this picture is the sharp divide in perceptions of the necessity of such distancing behaviors: We cannot expect people who believe these practices to be essential to respond emotionally to them in the same way as those who believe that they are a harmful over-reaction.⁴ The "expectations states" stream of social psychological theories posits that personally endorsing the recommendations of authorities widely regarded as legitimate in the polity should enhance wellbeing (e.g., Zelditch, 2018). Applying this to the epidemic suggests that people who endorsed distancing policies would have higher levels of wellbeing than others.

Just as the net effects of precautionary behaviors on wellbeing are ambiguous (with them likely increasing wellbeing through health protection and altruism and decreasing it through reduced social contact), the emerging quantitative research from the beginning of the pandemic is similarly mixed. A large study in New Zealand found no impact of their strict coronavirus lockdown on general wellbeing, although it did find a modest increase in mental distress (Sibley et al., 2020).⁵ On the other hand, a longitudinal study in Britain, found quite substantially worsened mental health in the wake of the pandemic – with a wellbeing drop attributable to the pandemic of 10% overall, and even worse for women (Banks & Xu, 2020). Intriguingly, research from Italy – which had one of the earliest severe outbreaks in Western Europe – suggested that emotional responses were substantially heterogeneous with different groups of people responding differently depending on a variety of factors including economic insecurity and gender (Cerbara et al., 2020). Unsurprisingly, in the US those who lost their jobs because of pandemic-related business closures were also not thriving, with increased rates of mental disorders especially among those who felt particularly economically uncertain (Yao & Wu, 2021). None of these studies looked at the

⁴ Indeed, there is considerable controversy in the research community itself over the impact of different voluntary and government mandated precautions against COVID-19 on health, hospitalization and mortality (e.g. Cerqueti et al., 2022; Herby et al., 2022).

⁵ Another study reports a different result (Greyling et al., 2021b).

impact of distancing behaviors specifically – rather they looked at the impact of the pandemic a whole – but this does suggest that there is not necessarily a single, society-wide impact of social distancing norms. For example, Greyling et al. (2021a) found that in South Africa stay-at-home orders *increased* happiness while other lockdown policies (e.g., bans on alcohol sales) decreased it.

Research Questions

Attitudes and Subjective Wellbeing

While we are aware of no existing research exploring the direct connection of attitudes towards distancing behaviors and subjective wellbeing; however, the ‘expectation states’ theory suggests a positive relationship. This leads us to Research Question 1.

RQ1: Do people who support distancing policies have higher subjective wellbeing, net of other influences?

Note that RQ 1 and all our other research questions are “*ceteris paribus*” questions, i.e., they concern net relationships controlling for other influences.

Precautionary Behavior and Subjective Wellbeing

The literature reviewed above is deeply ambiguous about the impact of precautionary/distancing behavior on subjective wellbeing. On the one hand, distancing has sociotropic, public-good implications and involves agency to protect one’s own health which would be expected to lead to a positive relationship between distancing behavior and subjective wellbeing. On the other hand, distancing behavior involves social constraint rather than freedom, so this aspect of distancing ought to lead to a negative relationship between distancing behavior and subjective wellbeing. The observed relationship is equally ambiguous. This raises Research Question 2:

RQ2: Does practicing distancing/ COVID-19 precautionary behaviors have a net positive or negative relationship with subjective wellbeing?

Emotional Costs of Precautionary Behavior and Subjective Wellbeing

Furthermore, we do not necessarily expect the connections between social distancing and wellbeing to be homogeneous across individuals. Findings from Italy (Cerbara et al., 2020), suggest substantially heterogeneous effects of the pandemic on happiness. Looking to other substantive domains, research suggests that the intrinsic (possibly culturally specific) unpleasantness of an action leads one to do it less often (Kauschke et al., 2019) and that this need not be mediated by changing attitudes or perceptions, but rather that emotion can directly influence behavior (Mills & Kleinman, 1988). It seems likely, therefore, that some people will find the COVID precautions more emotionally costly than others and that this

additional cost will reduce subjective wellbeing.⁶ Intuitively, we can all imagine the extroverted aspiring chef who, finding the lack of dinner parties crushing, is much less happy during the epidemic and the introverted hiker who, feeling little loss from the absence of social events, is quite as happy as before. This leads to Research Question 3:

RQ3: Do people who find social distancing behaviors emotionally costly have lower subjective wellbeing, all else equal?

Do Attitudes Moderate the Connection of Distancing Behavior and Wellbeing?

Evidence suggests also the people may be happier when their values align with the values of those around them (Sagiv & Schwartz, 2000) and the values promulgated by institutions they endorse as legitimate (Gibson et al., 2005; Zelditch, 2018). People may conform to norms and policies they do not support because they perceive the majority of others in their locality endorsing those norms and policies (Ridgeway, 2018), but this has long been known to generate dissonance (Brown et al., 1986) and so is likely to reduce wellbeing. This suggests an extension of the model that allows the relationship of practicing precautionary behaviors with subjective wellbeing to vary according to one's attitude toward precautionary behaviors (a multiplicative interaction). This leads to RQ 4:

RQ4: Does the match between distancing attitudes and distancing behavior moderate the net relationship of COVID-precautionary behavior and wellbeing?

Approaching the Research Questions

This paper aims to discover the separate (net) relationships of wellbeing to distancing behavior, to attitudes about distancing policies, and to the emotional costs of distancing, aside from their incidental connections with a host of other variables. In short, we focus on their net effects. To measure these as closely as possible our models control for objective and subjective background influences: demography and family background during primary socialization, current social situation including socioeconomic status, and perceived seriousness of COVID-19.

It is, of course, theoretically possible that people who enjoy high subjective wellbeing are thereby emotionally fortified and so experience lower emotional costs of precautionary behaviors. We provide some first evidence against this causal path

⁶ Note that positing a net statistical association of emotional costs with subjective wellbeing in no way assumes that this is the only direct effect of COVID-19 on subjective wellbeing. It does claim that emotional costs have a separate relationship to subjective wellbeing, independent of other variables. There are many other ways besides emotional cost in which distancing may affect individuals differently. We control for many of them – including age, gender, race and ethnicity, education, gender, occupation, income, and perception of the seriousness of COVID-19. Many others are also possible, but unless they are highly correlated with emotional cost of distancing, their omission will not bias our estimate of the effect of emotional costs.

with a structural equation model allowing reciprocal causation (Supplementary Material Appendix: RECIPROCAL EFFECTS).

Data: International Social Science Survey

International Social Science Survey Round 20, USA 2016–2019

Comparison data from pre-COVID times are from the International Social Science Survey Round 20, USA 2016–2019. The main wave (January 2019) is a representative US national sample ($N=1,778$) from Survey Sampling International's well-regarded internet panel (SSI; www.surveysampling.com). There were also two developmental waves ($N=1,323$ in 2016 and $N=1,173$ in 2017); for these issues analysis shows that they are indistinguishable from the main wave, so they are included.

International Social Science Survey Round 21, USA 2020

Most of analyses are based on the International Social Science Survey (ISSS) Round 21 with nine national surveys focusing on coronavirus and income inequality ready for analysis (so far). They began in April 2020 and have continued more or less monthly thereafter. Ten surveys, the last stretching into November 2020, are available for this analysis. $N=10,802$. The number of cases differs slightly between analyses as not all questions were asked on all the surveys.

All of the surveys are Amazon 'Mechanical Turk' internet-based samples. These can be conducted without risk of contracting coronavirus and, moreover, are swift, convenient, and cost effective. Internet samples of this sort have been widely used during the pandemic (e.g. Cornesse & Blom, 2020; Graham et al., 2020; McFadden et al., 2020) including to study the impact of the pandemic on wellbeing specifically (e.g., Ikeda et al., 2021; Lu et al., 2021).

Weighting

Extensive research shows that Mechanical Turk samples have proven remarkably accurate for multivariate analyses (Clifford et al., 2015; Springer et al., 2016; Weinberg et al., 2014), specifically for political affiliations and political attitudes (Clifford et al., 2015) and for science-related issues (Evans & Kelley, 2011). We have also found the same result when comparing our 2017–2019 internet samples to face-to-face samples in the US and many other countries (from the International Social Survey Programme's highly regarded "Inequality" series).

A disadvantage is that estimation of population means in internet samples may need to be adjusted because they typically over-represent younger, highly educated, and less religious people (e.g. Adida et al., 2021; Cornesse & Blom, 2020; Cornesse et al., 2020; McFadden et al., 2020), although correlations and regressions are very similar to those found in high-quality probability samples (Clifford et al., 2015;

Weinberg et al., 2014). Moreover, estimates from Pew's carefully curated internet sample and the ISSS mTurk samples match closely on many COVID-related questions, such as the extent of the decline in vaccine confidence over the period covered by this paper (Pew, 2020).

To be on the conservative side of the weighting issue (Bollen et al., 2016), we weight the sample according to the joint distribution of age and education (benchmark US Census Bureau) and then further adjust it according to childhood religious practice (attendance at public worship services in the family of origin (benchmark International Social Survey Programme Religion-3). Supplementary Material Appendix WEIGHTING & ROLE-PLAYING BIAS covers these matters in detail.⁷

Role-Playing Respondents Omitted

Another serious problem, not previously noted in the literature, is role-playing. Some respondents "Almost always gave their OWN TRUE VIEWS" rather than (as we put it in our question on the matter) "pretending to be someone else, playing a role instead of giving their own true views". In recent years a frightening proportion role-play, at least sometimes – perhaps 25% or 30% of all respondents.

Moreover, this is not a problem only, or even mostly, with Mechanical Turk samples but appears much more broadly. It seems to have emerged only in recent years in Mechanical Turk samples but was to our knowledge already present years earlier in a high-quality commercial panel (Kelley et al., 2021). We consider the matter in detail in Supplementary Material Appendix WEIGHTING & ROLE-PLAYING BIAS.

For this paper, we address the role-playing problem conservatively, dropping all respondents who admit to even occasionally role-playing.

Methods and Measurement

Methods

This paper uses descriptive statistics, correlations, factor analysis, and ordinary least squares regression analysis as well as structural equation methods. See also the discussion of our weighting approach above in the subsection "Weighting" of the

⁷ Despite the growing body of evidence here, it remains theoretically possible that if there are large interactions – differing effects in different groups – that are omitted from the model, the coefficient estimates will be incorrect. The long-established general theoretical result that multivariate models are robust to moderately large flaws in sampling representativeness only holds if the model is correct in functional form and in inclusion or exclusion of interactions (Berk 1983, Winship & Radbill, 1994; Bollen et al., 2016; a good recent summary is in Cornesse et al., 2020). We think this is unlikely to be a problem in our analysis, because (1) our check for functional form did not reveal any curvilinear effects that are robust across the different models and (2) prior research has not demonstrated robust interactions involving any of our control variables. Accordingly, it is reasonable provisionally to take the specification of the models as correct. Note that issues of incorrect specification are by no means unique to analyses based on non-probability samples, but need to be addressed in every parametric analysis.

"Data" section and in Supplementary Material Appendix WEIGHTING & ROLE-PLAYING BIAS.

Where possible, we use multiple-item scales to reduce random measurement error, thereby enhancing the precision of measurement of the statistical relationships of interest (e.g. Bryman, 2016).

We use correlations, factor analyses, and structural equation methods to evaluate the measurement properties of the scales using the "classical measurement model": judging the quality of a scale by several indications of whether the proposed items (measured variables) tap a single underlying latent (unmeasured) variable (Details are in Supplementary Material Appendix CLASSICAL MEASUREMENT MODEL). The criteria are: (1) high correlations among the candidate items; (2) in a factor analysis of the candidate items, they should all have high loadings on the same factor and no substantial cross-loadings on other factors; (3) all the candidate items should have parallel correlations with "criterion variables", variables outside the scale, e.g. one criterion variable might have very low correlations with all the candidate items, another might have very high correlations with all the criterion variables, so long as all are similar it does not matter whether they are high or low, positive or negative. However, if one candidate item has a very different correlation from the other candidate items with a particular criterion variable, that would be evidence that the candidate item with the deviant correlation might not belong in the scale (e.g., Bollen, 1989; Treiman, 2009a.) We supplement these analyses with Cronbach's Alpha which is sometimes called internal consistency alpha or internal consistency reliability (e.g., Bryman, 2016). We deal with these issues in detail in Supplementary Material Appendix CLASSICAL MEASUREMENT MODEL.

For our basic regression models of subjective wellbeing, we report robust standard errors. In our models for the most part we do not allow reciprocal causation (sometimes called reverse causation). This means that the regression coefficients we present are upper bounds on the coefficients we would find if reciprocal causation were allowed (Evans & Kelley, 2008) and hence are first approximations to causal effects. Persuasive models of reciprocal causation require strong assumptions and are often controversial in their specification, but we have estimated such a model as one angle on the causation issue (Supplementary Material Appendix RECIPROCAL EFFECTS). The findings indicate effects of emotional costs on SWB, but no reciprocal causation from SWB to emotional costs. This suggests that we can take the upper-bound estimates from the "plain vanilla" regression as likely to be accurate central estimates for the population regression coefficients we would find if reciprocal causation were allowed everywhere. Panel analysis currently in the data collection stage will prove fruitful here.

We focus on standardized regression coefficients because these allow sensible comparison between predictors measured in different metrics (for example dollars of income, years of education, or strength of agreement). Standardized coefficients reflect the strength of a variable's net relationships (often called "effects", in some disciplines called "beta coefficients") with the response/ dependent variable. Accordingly, we here focus on the standardized regression coefficients that estimate the strength or "tightness" of the relationship of subjective wellbeing to distancing behavior, attitudes towards distancing, and emotional costs

of distancing net of the other relationships in the model (Sirkin, 2006; Treiman, 2009a, b; Vittinghoff et al., 2005; Warne, 2020). Standardized regression coefficients give the change in the dependent variable (in standard deviation units) associated with a one-standard-deviation change in the predictor, net of the influences of the other predictors in the model. As such, they represent the relative influence of all the predictors in the model mapped onto a common scale so that their strength can readily be compared even when they are measured in different units. They do not reflect the steepness/magnitude of the relationship (the number of units by which Y changes, on average, with a one-unit change in X) which is revealed by the metric regression coefficient. Instead, the standardized coefficients reveal how tightly the values of the dependent/response variable gather around the regression line representing the linear relationship between variable X and variable Y, adjusting for the influence of other variables. They are akin to partial correlation coefficients, except that they mete out the shared contribution of the predictors so as to maximize the model's fit/ explanatory power.

They range between -1.0 (strongest possible negative net association) to 0 (no net association at all) to 1 (strongest possible positive net association). For individual-level variables, a common rule of thumb, expressed in absolute values is: < 0.05 = too weak/loose to matter; 0.05 – 0.09 = weak/loose, but non-ignorable, 0.10 – 0.19 = moderately strong/tight, 0.20 – 0.30 = strong, > 0.30 = very strong. They can be obtained either by z-scoring the variables and then estimating the regression equation in the usual way or alternatively by transforming the metric regression coefficients by a little algebra (Treiman, 2009a, b; Warne, 2020). T-statistics and p-values are mathematically identical for standardized and metric regression coefficients and metric coefficients differ only by a simple scale factor.

Item-missing data are rare in this dataset, but the surveys evolved as the epidemic proceeded so some questions were not asked in the earlier surveys. In particular, the question about the emotional cost of refraining from partying question was not asked in the first two surveys and is therefore missing completely at random (MCAR) for the respondents to those surveys. In order to be able to include that item in the intended composite scale using "prorating" (also called "available items imputation") the item needs to meet 3 standards (Chen et al., 2020; Mazza et al., 2015). (1) The item's mean (when present) is within the range of the means of the other candidate items for the scale. (2) The item's interitem correlations with the other candidate items for the scale are within the ranges of correlations among the other candidate items. (3) The correlations of the item with criterion variables are within the range of the correlations of the other candidate items with each of the criterion variables. These standards strongly echo the established criteria for the classical measurement model (described above). All these standards are met in the case of the emotional cost of refraining from partying, as will be seen below in the section on "Focal independent/predictor variables". Hence, when building the emotional costs scale, we therefore impute the scale score for the emotional cost of partying based on the questions that were asked, i.e. we use prorating (Supplementary Material Appendix EMOTIONAL COSTS).

In addition to the emotional cost items focusing on social deprivation, the surveys also included measurements of the emotional costs of "public" distancing behavior

(also called "civic" or "casual" distancing behavior such as standing 6 feet apart, wearing masks, and refraining from shaking hands. These items would normally be discarded on measurement-property grounds, but they are of considerable substantive and policy interest. Suffice it to say here that our analysis gives a good estimate of the *combined* effect of costs associated with both "social" and "public" distancing (Table 11, line 3) but that the split between them (Table 10, Column 5, last two lines) is uncertain. This issue is covered in detail in Supplementary Material Appendix EMOTIONAL COSTS.

Measurement Approach

Scoring Likert-type Items for Scaling

For the multiple-item scales, items were scored in equal intervals ('Likert' scoring) from a low of zero to a high of 100. For example, for 5 category answers: Strongly in favor (100), In favor (75), Mixed feelings, not sure (50), Opposed (25), and Strongly opposed (0).

This zero-to-one hundred scaling is for clarity. Any other equal interval scoring (for example, the original Likert 1–2–3–4–5) would produce mathematically identical standardized coefficients (also correlations, factor loadings, R-squares, etc.) and metric coefficients differing only by a simple shift factor: is essentially cosmetic. It is often convenient to score the predictor variables from 0 to 1 (0, 0.25, 0.50, 0.75, 1) and the response variables from 0 to 100 (0, 25, 50, 75, 100) because the metric regression coefficients then do not have leading zeroes. Research shows that, as Likert originally intended (1932), in diverse content domains respondents treat these as quantitative variables measured at equal intervals and that predicted values from regressions using equal interval scorings and scorings based on probit or logit analysis tend to correlate at about 0.98 or higher (e.g. Evans & Kelley, 2004; Ravallion et al., 2016). Accordingly, the Ockham's Razor criterion of preferring simplicity over complexity implies that researchers should use the equal-interval scoring for the Likert items they use in building multiple-item scales, so we follow that "intervalist" practice here.

Scale scores for each individual were computed as the unweighted mean of their values on the component items, thus giving each item equal logical weight. This is conceptually clearer than using a factor score; in practice, factor-score results hardly differ (results not shown but available on request).

If a respondent did not answer an item, it was omitted from their average, so scale scores are missing only for the (very few) respondents who failed to answer *any* of the component items in a scale.

Latent Variables and Scale Development

The scales for the different features of people's feelings about distancing, distancing policy attitudes, and distancing behaviors were optimized separately, because there is no reason to expect that their connections will be the same across domains. For example, all the behaviors might form one coherent package of

social practices, but people's feelings about the behaviors might clump together in quite distinct groups (for instance, with some people minding only the restrictions on contact with close friends and family while other find all social restrictions painful). Forming each scale independently minimizes random measurement error and hence achieves the most precise estimates of effects. Nonetheless, one might be curious about whether the results would be different if we used the same maximum set of indicators for each domain (as many of the behaviors as were present in all the domains) or if we used the same minimum set of indicators for each domain (only the indicators that met classical measurement criteria standards for the emotional cost of distancing scale). The results of these alternatives are not substantially different (details not shown but available on request).

Descriptive Statistics

For all the variables in the study, the means and standard deviations are in in Supplementary Material Appendix SAMPLE DESCRIPTION. The full set of correlations is in Supplementary Material Appendix CORRELATIONS. Supplementary Material Appendix MISSING DATA gives detailed information on missing data.

Dependent/ Response Variable

Wellbeing: Life Satisfaction, Happiness

We measure wellbeing with two familiar, widely used questions, following the wording from the European Quality of Life surveys. The answer choices are provided on a 1 to 10 range with the endpoints anchored with words. The first question asked "All things considered, HOW SATISFIED would you say you are with your life these days? On a scale of 1 to 10 are you: ..." with answer categories "1 (Very dissatisfied)" 2, 3, 4, 5, 6, 7, 8, 9, "10 (Very SATISFIED)". The histogram below shows the distribution of responses (Fig. 1).

The second question asked "Taking all things together on a scale of 1 to 10, HOW HAPPY would you say YOU are:" with answers similarly ranging from "1 (Very unhappy)" to "10 (Very HAPPY)". The distribution of the responses is in Fig. 2.

There would be less risk of random measurement error if these were measured as 0 to 10 rather than 1 to 10,⁸ but respondents nonetheless seem to have little trouble answering them.

⁸ 0 to 10 is better because, even with the clearest instructions, some respondents get the 1 to 10 scale backwards (treating 1 as the highest score, perhaps by analogy to first place in a race) but virtually nobody ever gets the 0 to 10 scale backwards. This issue is diminished here by the inclusion of verbal labels at the endpoints. The 0 to 10 scale is also better because it has an explicit midpoint which is important for some topics (but not for wellbeing). Scales on a 1 to 10 range may have some distortion, because many respondents apparently think 5 is the midpoint.

Note that for clarity, the answer options for both subjective wellbeing items are rescored as points out of 100. Thus, answer category "1 (Very unhappy)" is scored zero and category "10 (Very HAPPY)" is scored 100 and all the others scored at equal intervals in-between.⁹ This is a purely cosmetic change which makes results more intuitive. Thus, for example, satisfaction among higher professionals (doctors, professors, etc.) is 65 points out of 100, compared to 56 points for unskilled workers.

Our subjective wellbeing scale (Table 1) is made by averaging these two items. It fits the classical measurement model well (Supplementary Material Appendix CLASSICAL MEASUREMENT MODEL). Cronbach's alpha (a standard measure of reliability) is high: 0.88.

Interitem Correlation The two subjective wellbeing items have a very strong correlation of 0.77 (Table 1: Panel labeled "Interitem correlations of candidate variables"). That meets the first requirement of the classical measurement model for strong interitem correlations.¹⁰ The means are near 63 points out of 100 with standard deviations around 25 (Table 1: Panel labeled, "Descriptive statistics").

Confirmatory Factor Analysis The two subjective wellbeing item both have high factor loadings of 0.88 and 0.85 respectively.

Correlations with Criterion Variables Next, consider the correlations of the two subjective wellbeing items with criterion variables which will help us assess whether the two wellbeing items behave similarly (Table 1, Panel labeled "Correlations of criterion variables with subjective wellbeing items"). First, the correlations with gender are both near zero: The correlation of "Male" with "Satisfied" is 0.01 and the correlation of "Male" with "Happy" is 0.01. They are very similar. Reading across each of the rest of the rows will also reveal closely similar correlations of each criterion variable with both candidate items. For example, "Religious belief" correlates with both candidate items at around 0.21. Unemployment correlates with both candidate items at around -0.10. Political party preference

⁹ Thus "10 (Very HAPPY)" is 100; 9 is 88.8; 8 is 77.7; 7 is 66.6; 6 is 55.5; 5 is 44.4; 4 is 33.3; 3 is 22.2; 2 is 11.1; "1 (Very unhappy)" is 0.

¹⁰ How high the correlations need to be to meet this criterion depends on how many items the scale has, so much lower correlations are acceptable among personality measurement candidate items because the scales can contain dozens of items. Early measurement in the social sciences where many items were dichotomous and hence larded with random measurement error tended to tolerate correlations as low as 0.3, although 0.4 would be a more common minimum today as questions with multiple answer categories (which have less random measurement error) have become more common. A stricter standard such as 0.5 is appropriate for very short scales, such as the 2-item scale considered here. With a correlation of .77 our two items easily meet even the stricter threshold. See Appendix CLASSICAL MEASUREMENT MODEL.

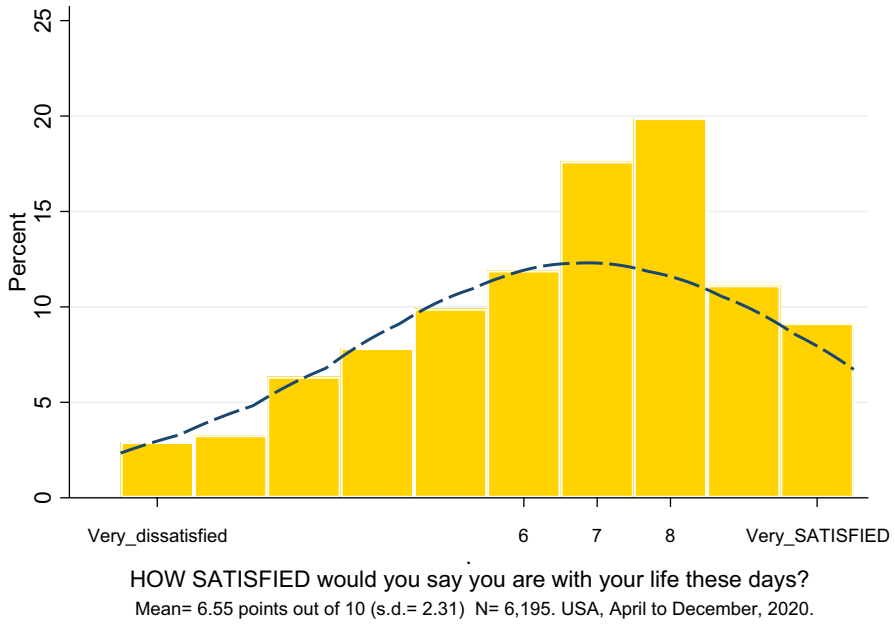


Fig. 1 Life satisfaction. Percent distribution (bars) and kernel density (line). $N=6,195$ US adults. April–November 2020



Fig. 2 Happiness. Percent distribution (bars) and kernel density (line). $N=6208$ American adults. April to November 2020

Table 1 Subjective wellbeing scale. Classical measurement assessment: Interitem correlations, CFA loadings, correlations with criterion variables. April to November 2020

		Correlations: Candidate items for subjective wellbeing scale	
		Satisfied?	Happy?
Interitem correlations of candidate variables	Satisfied? (0 to 100)	1,00	
	Happy? (0 to 100)	,77	1,00
Descriptive statistics	Mean (0 to 100)	62	64
	Standard deviation	26	25
Factor loading	CFA loading	,88	,85
Correlations of criterion variables with subjective wellbeing items	Male (0 or 1)	,01	,01
	Black (0 or 1)	,01	,02
	Hispanic (0 or 1)	,00	,02
	Age (years)	,05	,05
	Churchgoing at age 14 (ln # times per year)	,08	,09
	Religious belief (0 to 100)	,21	,22
	Education (years)	,13	,09
	Married (1 or 0)	,21	,19
	Family income (ln \$1000 s)	,27	,23
	Unemployed (0 or 1)	-,11	-,10
Democratic Party (0–1)	-,12	-,14	
Time (week of epidemic)	-,08	-,03	

[1] The pairwise number of cases ranges from 6165 to 6208 for the correlations shown here. Missing data was extremely rare and virtually random (Supplementary materials appendix MISSING DATA). The scale of subjective wellbeing used in subsequent analysis is the mean for each respondent on satisfaction and happiness where each item is scored in equal intervals from a low of zero to a high of 100. Sample weighted to reflect the joint distribution of age and education (benchmark US Census Bureau) and the univariate distribution of churchgoing at age 14 (benchmark ISSP Religion-3). Role-playing respondents are omitted (Appendix WEIGHTING & ROLE-PLAYING)

(scored 0 for strongly preferring the Republican party to 1 for strongly preferring the Democratic Party)¹¹ has correlations near -0.13 for both candidate items (Democrats less happy.). In some instances, one candidate variable has a fractionally stronger correlation with the criterion variables than the other does, but these differences are all small enough to be ignorable and might well be due to chance. Thus, the two candidate items have similar patterns of correlations with criterion variables. This meets

¹¹ The party preference link to wellbeing has a weaker absolute value correlation with wellbeing than does the self-identified left–right ideology (not shown), but party preference is a better control variable here (helping us to get close to "pure" estimates of the effects of interest), because, throughout the epidemic, COVID behavior and attitudes have been more closely connected with party than with ideology (these, of course, are correlated, but far from identical).

Table 2 Frequency of attitude toward distancing policies (strong disapproval = 0, strong approval = 100, other answers at equal intervals in between)

	Frequency of Answer (Percent)				
	Strongly In Favor	In Favor	Mixed feelings, not sure	Oppose	Strongly Oppose
Requiring to stand six feet apart	57	26	10	5	2
Not shake hands	62	23	9	4	2
Wash hands often	71	19	6	2	2
No parties, bars, etc	55	23	12	6	4
Work from home, if possible	49	29	13	6	3

Sample weighted to reflect the joint distribution of age and education (benchmark US Current Population Survey) and the univariate distribution of childhood churchgoing (benchmark ISSP Religion-3). Role-playing respondents are omitted (Supplementary Material Appendix WEIGHTING & ROLE-PLAYING BIAS). USA April to November 2020

the "similar correlations with criterion variables" criterion of the classical measurement model. We construct a scale as a simple average of the two items (giving them equal weight).

This close parallelism of correlations with criterion variables accords with the requirement of the classical measurement model for the candidate items with similar factor loadings (as here) also having similar correlations with criterion variables.

Bivariate Pearson correlations with other variables in the study are in Supplementary Material Appendix CLASSICAL MEASUREMENT MODEL. Details on missing data are in Supplementary Material Appendix MISSING DATA.

In the questionnaire, these two items have a fixed location in a module entitled "How are you doing?" in all ten surveys, so question order effects will not be a problem for this analysis.

Focal Independent/ Predictor Variables

Favorable or Unfavorable Attitude Toward Distancing (Scale Kek-3)

Our first research question concerns the statistical relationship of subjective wellbeing to policy attitudes about the group of COVID-preventative behaviors generally collectively known "social distancing". The introduction to these questions asks "Considering the WAY THINGS ARE THIS WEEK HERE IN THIS TOWN – and balancing the good with the bad – what do you think about these distancing policies for reducing coronavirus risks...". The five individual policy items, then, are: (1) Requiring people to stand least 6 feet away from strangers in public, (2) Not shaking hands when people meet, (3) Washing hands after going

Table 3 Measurement of attitude toward distancing policies (strong disapproval = 0, strong approval = 1), other answers at equal intervals in between). Sample weighted to reflect the joint distribution of age and education (benchmark US Current Population Survey) and the univariate distribution of childhood churchgoing (benchmark ISSP Religion-3). Role-playing respondents are omitted (Supplementary Material Appendix WEIGHTING & ROLE-PLAYING BIAS). Data are for the USA, April to November 2020. Cronbach’s Alpha = .87

Interitem correlations	Candidate items for scale of attitudes towards distancing policy. All are measured on a range of 0 (strongly disapprove) to 1 (strongly approve) at equal intervals				
	(a)	(b)	(c)	(d)	(e)
Approve: Stand six feet apart (a)	1.000				
Approve: Not shake hands (b)	0.700	1			
Approve: Wash hands often (c)	0.508	0.610	1		
Approve: No parties, bars, etc.(d)	0.663	0.674	0.513	1	
Approve: Work from home (e)	0.600	0.574	0.469	0.642	1
Descriptive statistics					
Mean	.83	.83	.87	.79	.79
Standard deviation	.24	.25	.22	.27	.26
Factor loading					
CFA loading	0.81	0.84	0.66	0.82	0.73
Correlations with criterion variables					
Male (0 or 1)	-0.082	-0.124	-0.155	-0.078	-0.083
Black (0 or 1)	0.057	-0.013	-0.041	0.020	0.012
Hispanic (0 or 1)	0.009	0.027	0.027	0.012	0.014
Age (years)	0.009	0.007	0.031	0.015	-0.029
Churchgoing at age 14 (ln # times per year)	0.025	0.049	0.085	0.041	-0.002
Religious belief (0 to 100)	-0.058	-0.102	-0.050	-0.076	-0.103
Education (years)	0.029	0.023	-0.041	0.022	0.017
Married (1 or 0)	-0.012	-0.044	-0.053	-0.027	-0.014
Family income (ln \$1000 s)	0.024	-0.032	-0.045	-0.035	0.019
Unemployed (0 or 1)	0.012	0.014	0.017	0.011	0.011
Democratic Party (0–1)	0.224	0.210	0.142	0.242	0.211
Time (week of epidemic)	-0.064	-0.102	-0.075	-0.086	-0.097
Correlations with dependent variable, subjective wellbeing	0.086	0.063	0.068	0.048	0.058

*The pairwise numbers of cases range from 6105 to 6223 for the correlations shown here. Missing data were extremely rare and virtually random (Supplementary Material Appendix MISSING DATA). The scale used in subsequent analysis is the mean for each respondent on the attitude items, each item rescored to range in equal intervals from 0 through 1

out in public, (4) Avoiding parties, bars, and restaurants, and (5) Requiring people to work from home if possible? Distributions are in Table 2.

The balance of opinion is clearly favorable for all the questions, although the full spectrum of opinion is also present for each question. Least popular is working from

home, with 78% in favor (strongly or mildly). Most popular is frequent hand washing, with 90% in favor (strongly or mildly).

Interitem Correlations The candidate items for the distancing-policy-attitudes scale are strongly connected among themselves, with all their correlations being over 0.4 (Table 3, top panel, "Interitem correlations"). Thus, the classical measurement model's standard of strong correlations among the candidate items for the scale is met.

Descriptive Statistics Means on these items are all well towards the "Strongly approve" end, ranging from 79 to 87 points out of 100 (Table 3, Panel labeled "Descriptive statistics"). The standard deviations are in the neighborhood of 25.

Confirmatory Factor Analysis The various candidate items for the scale of attitudes towards social distancing policy have high loadings on a single central factor: 0.81 for standing 6 feet away, 0.84 for not shaking hands, 0.82 for avoiding bars and parties, and slightly lower but a still robust 0.66 for washing hands and 0.73 for working from home.

Correlations with Criterion Variables The candidate items for the scale of attitudes towards social distancing policy have similar/ parallel correlations with criterion variables (Table 3, Panel labeled "Correlations with criterion variables"). For example, all the candidate items have near-zero correlations with age, all have correlations in the range -0.05 to -0.10 with religious belief, and all have correlations in the range 0.14 to 0.24 with political party preference (strong Democrat is high). None of the candidate items has a strongly deviant correlation with any of the criterion variables. This pattern of results aligns with the similar/parallel correlations with criterion variables requirement of the classical measurement model.

Considering the candidate items as a multiple-item scale yields a strong Cronbach's Alpha of 0.87 (0.70 is the common criterion for acceptable internal coherence, 0.75 is a stricter criterion; the Cronbach's Alpha for this scale exceeds even the stricter criterion).

Taken together, these results (as well as face validity) justify adding all the candidate items together to build a scale of attitudes towards distancing.¹²

Distancing Behavior (Scale KEK-2)

The degree to which we can treat self-reports of distancing behavior as objective depends on their accuracy. Fortunately, preliminary findings on the accuracy of self-reports of distancing behavior suggest little bias (Gollwitzer et al., 2022; Jensen, 2020), so it is reasonable to take "mostly objective" as a working hypothesis for distancing self-reports.

¹² (This scale, KEK-3 **Attitude to Social Distancing**, is copyright © 2020 by the authors but may be freely used by other researchers so long as its origin is acknowledged by reference to this paper.)

Table 4 Frequency of practicing social distancing (never=0, always=100, other answers at equal intervals in between). Sample weighted to reflect the joint distribution of age and education (benchmark US Current Population Survey) and the univariate distribution of childhood churchgoing (benchmark ISSP Religion-3). Role-playing respondents are omitted (Supplementary Material Appendix WEIGHTING & ROLE-PLAYING BIAS). USA April to November 2020

	Frequency of Answer (Percent)						
	Always	Almost Every Time	Most of the Time	Half of the time	Sometimes	Occasionally	Never
Standing six feet apart	38	34	16	4	5	2	1
Not shake hands	53	15	11	4	4	2	2
Wash hands often	63	18	9	4	4	2	1
No parties, bars, etc	51	19	15	5	6	3	3
Avoiding going to other people houses	43	21	15	6	7	4	5

Questions and Distributions The questions on distancing behavior were introduced with the stem “Some precautions help reduce the risk that YOU CATCH coronavirus...” and included five question: (1) “One precaution is to stand at least 6 feet away from strangers in public. Of course not everyone does that. But what about you, yourself – do YOU stand 6 feet away?”, (2) “Another precaution is not shaking hands when you meet someone (perhaps touching elbows instead, or just saying hello). Do you avoid shaking hands?”, (3) “How about washing your hands after you have been out in public?”, (4) “How about avoiding parties, stores, and restaurants?”, (5) “Avoiding going to other people’s houses or having them come to your house?”. Table 4 gives the percent distributions.

As part of the same module of items, we also asked about working from home (one question), and self-quarantining (two questions). Analysis showed that, in the minds of respondents, the working-from-home item and the self-quarantining items raised several somewhat distinct issues from the other distancing items, and so do not fit the classical measurement model for a combined unidimensional scale. They are therefore not included.

A question on wearing a mask was also asked in the same module (and two similar questions on masks asked elsewhere in the survey). They raise more difficult issues, with good arguments both for and against including them in the scale.¹³ In practice, including the mask questions makes no noticeable difference here (results not shown) and so we omit them. An extensive analysis is given elsewhere (Evans et al., 2021).

The measurement analysis of candidates for the distancing-behavior scale is in Table 5.

¹³ Among other issues, the implications of mask wearing changed over the course of the epidemic as the CDC had originally recommended against this practice before updating their recommendation.

Table 5 Measurement of distancing behavior (never=0, always=100, other answers at equal intervals in between). Sample weighted to reflect the joint distribution of age and education (benchmark US Current Population Survey) and the univariate distribution of childhood churchgoing (benchmark ISSP Religion-3). Role-playing respondents are omitted (Supplementary Material Appendix WEIGHTING & ROLE-PLAYING BIAS). Data are for the USA April to November 2020. Cronbach's Alpha = .80

Interitem correlations	Candidate items for scale of distancing behavior. All are measured on a range of 0 (never) to 100 (always) at equal intervals				
	(a)	(b)	(c)	(d)	(e)
Behavior: Stand six feet apart (a)	1.000				
Behavior: Not shake hands (b)	0.499	1			
Behavior: Wash hands often (c)	0.471	0.508	1		
Behavior: No parties, bars, etc.(d)	0.477	0.472	0.513	1	
Behavior: No home visits (e)	0.413	0.429	0.388	0.5635	1
Descriptive statistics					
Mean	81	87	87	81	76
Standard deviation	21	24	21	26	29
Factor loading					
CFA loading	0.7	0.7	0.71	0.72	0.68
Correlations with criterion variables					
Male (0 or 1)	-0.059	-0.143	-0.141	-0.019	-0.063
Black (0 or 1)	-0.029	-0.083	-0.043	-0.008	-0.015
Hispanic (0 or 1)	0.010	0.009	0.032	0.009	0.003
Age (years)	0.059	0.064	0.085	0.042	0.103
Churchgoing at age 14 (ln # times per year)	0.017	0.062	0.059	0.030	0.031
Religious belief (0 to 100)	-0.002	-0.104	0.006	-0.046	-0.079
Education (years)	0.008	-0.034	-0.039	0.000	0.030
Married (1 or 0)	-0.001	-0.063	-0.001	0.006	-0.010
Family income (ln \$1000 s)	0.030	-0.058	-0.007	-0.030	-0.010
Unemployed (0 or 1)	0.015	0.022	0.020	-0.015	0.002
Democratic Party (0–1)	0.130	0.153	0.106	0.151	0.162
Time (week of epidemic)	-0.066	-0.102	-0.102	-0.138	-0.208
Correlations with dependent variable, subjective wellbeing	0.118	0.031	0.074	0.040	0.013

The item on partying was not asked in early surveys; for it the pairwise number of cases range from 3402 to 3412 for the correlations shown here. For the other items the pairwise number of cases range from 6060 to 6223. Missing data on questions that were asked was extremely rare and virtually random (Supplementary Material Appendix MISSING DATA). The scale used in subsequent analysis is the mean for each respondent of the emotional costs questions asked (each rescored to range in equal intervals from 0 through 100). For the early surveys where the partying question was not asked and is therefore missing completely at random (MCAR), this imputes the scale score on the basis of the questions that were asked, a procedure known in some disciplines as prorating or available item imputation, which is justified if the candidate items (including the one with the missing data) otherwise fit the classical measurement model (Chen et al., 2020; Mazza et al., 2015)

Interitem Correlations Are mostly in the 0.4 to 0.5 range with one fractionally lower (0.39) for refraining from home visiting with frequency of washing hands and with one somewhat higher (0.56) for refraining from home visiting and refraining from partygoing. Cronbach's Alpha is 0.80, showing good internal consistency.

Descriptive Statistics The means of the candidate items are similar and high, ranging from 76 points out of 100 (refraining from home visiting) to 87 (washing hands and refraining from shaking hands). The standard deviations all fall with the range of 20 to 30.

Confirmatory Factor Analysis The various candidate items for the scale of social distancing practices have factor loadings in the 0.68 to 0.72 range.

Correlations with Criterion Variables Most of the correlations with the criterion variables are very small. Since our criterion variable will be other predictors in our model of subjective wellbeing. This is an indication that there will be little collinearity involving the distancing behavior scale. On the whole, all the candidate variables have correlations with each criterion variables that are very similar. We can summarize by taking each candidate item's average deviation from the mean correlation for all the candidate items and then averaging that across all 12 criterion variables. These average deviations are near zero for all the candidate items: All are below 0.02. This suggests strong similarity/ parallel patterns of correlations of the candidate items with the criterion variables.

These findings justify creating a multiple-item scale of distancing behavior from these five items.

Emotional Cost of Social Distancing (scale KEK-6)

Many of the precautions we take against coronavirus have intrinsic disadvantages (and occasionally advantages for some people). Our surveys asked about a range of activities that have been publicly discussed as precautionary measures endorsed by public health authorities and by governments, focusing specifically on how respondents "PERSONALLY FEEL" (emphasis in the question) about taking these precautions. Note that the question is about how the respondent themselves feels about taking these precautions, not about whether they think the precautions are effective, or whether they actually do them, or whether they think they are burdensome to others. These include issues involving sociability and social deprivation involving restriction of social opportunities (dining out, parties, bars, visiting, overnighing with friends, and having friends overnight with you). Many people find these activities delightful, but some find them onerous and welcome a socially acceptable excuse to avoid them. Although we specifically asked respondents to report their feelings about these precautions apart from the effect of the precautions' effects on the risk of infection, it is possible that some respondents are including a moral glow from

Table 6 Emotional cost of distancing policies (terrible=0, excellent=100, other answers at equal intervals in between). Sample weighted to reflect the joint distribution of age and education (benchmark US Current Population Survey) and the univariate distribution of childhood churchgoing (benchmark ISSP Religion-3). Role-playing respondents are omitted (Supplementary Material Appendix EMOTIONAL COSTS). USA April to November 2020

	Frequency of Answer (Percent)						
	Terrible	Very Bad	Bad	Neutral	Good	Very Good	Excellent
Not going to restaurants	9	11	24	22	14	11	10
Not going to parties, bars, theater	7	10	18	22	14	12	17
Not visiting friends homes	13	15	24	20	12	9	8
Not stay overnight at friend's place	9	10	17	32	11	10	9
Not having friends visit your home	12	12	23	21	12	10	10
Not have a friend stay overnight	9	10	18	29	12	10	12

conforming to authorities' recommendations or from an altruistic feeling of protecting others in their reports of the emotional costs of the various behaviors.

We focus on social deprivation it has formed the focus of public discussion and public health interest in the impact of the epidemic on wellbeing and mental health. But we also provide a broader perspective on the potential emotional costs of wellbeing in Supplementary Material Appendix EMOTIONAL COSTS. We also asked about issues involving wearing masks when out in public; and issues involving public/ civic behaviors (staying 6 feet apart, not shaking hands, washing hands frequently). Feelings about all of them are correlated in somewhat complex ways but the core is social deprivation. For this paper we focus on that. Details are in the EMOTIONAL COSTS Appendix.

The answer categories are: Terrible- Very Bad – Bad- Neutral – Good-Very Good – Excellent. They are modeled on Andrews and Withey's (1976) well-known formulation of feelings about different aspects of life, in use since the 1970s and with over 8,000 citations. The frequencies are in Table 6.

As noted earlier (in section Theory and Prior Research), prior literature hints at the possibility that the experience of distancing has very different emotional costs to different people, but does not explore this in detail.¹⁴ The distributions above in Table 6 support that view – people's feelings about each of the distancing actions cover the full spectrum of feelings from those who feel "terrible" about the action to those who feel it is "excellent", perhaps enjoying a socially acceptable respite from onerous obligations or taking pleasure in performing a sociotropic act.

Results of the analysis of the measurement properties of emotional costs are in Table 7. Do these items all tap the same underlying concept, the same latent variable? To find out, we begin by perusing the interitem correlations (Table 7).

¹⁴ (This scale, KEK-6 **Social Distancing Deprivation**, is copyright © 2020 by the authors but may be freely used by other researchers so long as its origin is acknowledged by reference to this paper.)

Table 7 Measurement of emotional costs of distancing behavior (Excellent=0, Terrible=100, other answers at equal intervals in between). Sample weighted to reflect the joint distribution of age and education (benchmark US Current Population Survey) and the univariate distribution of childhood churchgoing (benchmark ISSP Religion-3). Role-playing respondents are omitted (Supplementary Material Appendix WEIGHTING & ROLE-PLAYING BIAS). Data are for the USA April to November 2020. Cronbach's alpha = .94

Interitem correlations: Emotional cost of...	Candidate items for scale of emotional costs. All are measured on a range of 0 (excellent) to 100 (terrible) at equal intervals							
	(a)	(b)	(c)	(d)	(e)	(f)	(g)	(h)
Not eating out at restaurants (a)	1.00							
No parties, bars, theaters (b)	0.74	1.00						
Not visit friends at home (c)	0.70	0.69	1.00					
Not overnight with friends (d)	0.67	0.61	0.79	1.00				
No friends visit you (e)	0.65	0.62	0.83	0.78	1.00			
No overnight visitors (f)	0.65	0.60	0.76	0.82	0.83	1.00		
Work from home (g)	0.46	0.45	0.50	0.48	0.51	0.51	1.00	
Strict isolation at home (h)	0.59	0.61	0.69	0.60	0.66	0.61	0.52	1.00
Descriptive statistics								
Mean	45	51	56	50	53	49	45	56
Standard deviation	30	29	29	28	29	29	29	32
Factor loading								
Exploratory loading	0.76	0.79	0.9	0.86	0.89	0.87	0.58	0.74
CFA loading	0.75	0.78	0.91	0.87	0.91	0.87	0.59	0.75
Correlations with criterion variables								
Male (0 or 1)	-0.09	-0.02	-0.12	-0.07	-0.09	-0.09	0.02	-0.11
Black (0 or 1)	-0.12	-0.09	-0.14	-0.10	-0.13	-0.11	-0.07	-0.11
Hispanic (0 or 1)	-0.02	-0.01	-0.02	-0.01	-0.01	-0.01	0.01	-0.01
Age (years)	0.05	-0.01	0.05	0.02	0.04	0.02	0.01	0.08
Churchgoing at age 14 (ln # per year)	-0.01	-0.02	0.02	0.00	0.03	0.01	0.00	0.03
Religious belief (0 to 100)	-0.07	-0.05	-0.07	-0.04	-0.05	-0.02	-0.01	-0.04
Education (years)	-0.11	-0.01	-0.11	-0.08	-0.10	-0.08	-0.04	-0.07
Married (1 or 0)	-0.04	-0.01	-0.05	-0.08	-0.03	-0.05	-0.02	-0.06
Family income (ln \$1000 s)	-0.04	0.02	-0.05	-0.05	-0.02	-0.02	-0.03	-0.04
Unemployed (0 or 1)	-0.01	0.00	-0.01	0.01	0.00	0.00	0.05	-0.01
Democratic Party (0-1)	-0.03	0.01	0.02	0.02	0.01	0.01	-0.01	-0.01
Time (week of epidemic)	0.01	0.04	-0.01	0.00	0.00	0.03	0.02	0.05
Correlations with dependent variable, subjective wellbeing	-0.12	-0.11	-0.09	-0.12	-0.08	-0.09	-0.09	-0.11

The item on partying was not asked in early surveys; for it the pairwise number of cases range from 3402 to 3412 for the correlations shown here. For the other items the pairwise number of cases range from 6060 to 6223. Missing data on questions that were asked was extremely rare and virtually random (Supplementary Material Appendix MISSING DATA). The scale used in subsequent analysis is the mean for each respondent of the emotional costs questions asked (each rescored to range in equal intervals from 0 through 1). For the early surveys where the partying question was not asked and is therefore missing completely at random (MCAR), this imputes the scale score on the basis of the questions that were asked, a procedure known in some disciplines as "prorating" or "available item imputation", which is justified if the candidate items (including the one with the missing data) otherwise fit the classical measurement model (Chen et al., 2020; Mazza et al., 2015)

Interitem Correlations The average correlation among the candidate items for the scale is 0.64. The correlations of the emotional costs of working from home with all the other items are well below the other interitem correlations, so one could argue that including this item does not strengthen the scale. Nevertheless, we retain the item both because its correlations are not objectively weak (just weaker than the others) and because of its substantive interest.

Descriptive Statistics The candidate items have similar descriptive statistics. The means range from 45 to 56, close to the neutral point of 50. The standard deviations are all in the range 28 to 32.

Confirmatory Factor Analysis The various candidate items for the scale of the emotional cost of social distancing practices generally have high factor loadings in confirmatory factor analysis, with factor loadings for all items excepting working from home having loadings between 0.75 and 0.91. Working from home has a slightly lower loading of 0.59.

Correlations of Criterion Variables With the candidate items are not strong, but their patterning conforms to the expectation of similarity/ parallelism. None of the candidate items has a correlation profile that is systematically substantially higher or lower than the others. To summarize, we can (1) compute the mean correlation of the candidate items with each criterion variable; then (2) compute the deviations for each candidate item from those means; and then (3) take the average deviation. If the average deviations for a candidate item were positive and strong that would be evidence against similarity. Equally, if the deviations were negative and strong, that would also be evidence against similarity. For these candidate items, those average deviations range between -0.01 and 0.02, all very close to zero. This evidence favors the view that the candidate items have similar correlations with criterion variables.

In sum, the candidate items meet the standards of the classical measurement model as measurements of a single latent variable, and we therefore combine them into a composite scale.

These actions are strongly concerned with social deprivation, especially of strong ties (or potentially strong) ties. But a great deal of the precautionary behavior mandated by authorities to slow the epidemic concerned modifying our public interpersonal behavior and social interactions, behaviors such as standing six feet apart in public, not shaking hands, and wearing masks for example. These too entail at least modest emotional costs (Figure &&&) but the situation here is analytically complex, because these public behaviors are neither clearly a part of the latent social deprivation variable tapped by the candidate items we have analyzed above, nor are they clearly separate from it. For the present, we have treated them separately in our main model in Table 10 and have assessed their joint importance with a sheaf coefficient (Whitt, 1986) in Table 11. Supplementary Material Appendix EMOTIONAL COSTS details the approach and procedures.

Table 8 Perceived severity of coronavirus for you and very close others (No problem=0, Terrible=100, other answers at equal intervals in between). Percent distributions of answers to candidate items for severity scale. USA April to November 2020

How bad would it be if...	Frequency of answer				
	Terrible (scored 100)	Very bad (75)	Bad (50)	Not too bad (25)	No prob- lem (0)
You got coronavirus?	27	24	24	16	9
If your spouse got it?	33	24	20	14	8
Your best friend?	24	24	27	16	9

[1] Details in Supplementary Material Appendix SEVERITY OF COVID. Sample weighted to reflect the joint distribution of age and education (benchmark US Current Population Survey) and the univariate distribution of childhood churchgoing (benchmark ISSP Religion-3). Role-playing respondents are omitted (Supplementary Material Appendix WEIGHTING & ROLE-PLAYING BIAS)

Perception of Severity: How Bad would it be to get Coronavirus?

Views about severity, about how bad it would be to get coronavirus, were measured straightforwardly: with three questions asking “How bad would it be ...” “If you got coronavirus?”, “If your spouse got it?”, “Your best friend?”. Answers ranged from “Terrible” (scored 100) to “Very Bad” (scored 50) to “Not Too Bad” (scored 25) to “No Problem” (Scored 0). Table 8 gives the frequency distributions.

The distributions above show that the balance of the general public’s perceptions of the severity of getting COVID-19 during April–November 2020 tilted towards the severe end, but nonetheless spanned the full spectrum.

For better or for worse from a public health point of view, these perceptions did not change systematically over time (see the correlations with criterion variables in Table 9).

All 3 candidate items for the perceived severity scale have high interitem correlations, have high factor loadings on a single factor, and all have similar correlations with criterion variables. Thus, they satisfy the requirements of the classical measurement model. Hence, we are justified in combining them into a multiple-item scale. For further details, see Supplementary Material Appendix: CLASSICAL MEASUREMENT MODEL.

Control Variables (Non-Focal Independent Variables)

Demographic and Socioeconomic Control Variables

Deep background controls include childhood/ primary socialization conditions and ascriptive/ demographic traits that have broad consequences for other people’s expectations about you. In keeping with Clark’s (2018) call for inclusion of primary socialization influences in models of SWB, we include parents’ education (years completed, average of the two, if only one available, we use that); parents’ occupational status (see Supplementary Material Appendix VERTICAL OCCUPATIONAL

Table 9 Measurement of perceived severity: How bad is getting coronavirus? Sample weighted to reflect the joint distribution of age and education (benchmark US Current Population Survey) and the univariate distribution of childhood churchgoing (benchmark ISSP Religion-3). Role-playing respondents are omitted (Supplementary Material Appendix WEIGHTING & ROLE-PLAYING BIAS). USA April to November 2020. Cronbach's alpha = .90

	Candidate items for scale of perceived severity of getting COVID-19. All items measured on a range of 0 (no problem) to 100 (terrible) at equal intervals		
	(a)	(b)	(c)
Interitem correlations			
How bad if you got it (a)	1.000		
How bad if spouse got it (b)	0.766	1	
How bad if best friend got it (c)	0.769	0.788	1
Descriptive statistics			
Mean	61	65	59
Standard deviation	32.1	32.5	31.5
Factor loading			
CFA loading	0.86	0.90	0.89
Correlations with criterion variables			
Male (0 or 1)	-0.061	-0.057	-0.088
Black (0 or 1)	-0.011	-0.043	-0.024
Hispanic (0 or 1)	0.001	-0.011	0.011
Age (years)	0.061	0.052	0.059
Churchgoing at age 14 (ln # times per year)	0.044	0.083	0.060
Religious belief (0 to 100)	0.041	0.032	0.034
Education (years)	-0.013	-0.010	-0.019
Married (1 or 0)	-0.008	0.057	-0.023
Family income (ln \$1000 s)	-0.059	-0.057	-0.059
Unemployed (0 or 1)	-0.022	0.009	-0.006
Democratic Party (0–1)	0.159	0.153	0.177
Time (week of epidemic)	-0.043	-0.062	-0.072
Correlations with dependent variable, subjective wellbeing	-0.068	-0.021	-0.048

*The pairwise number of cases range from 4778 to 6223 for the correlations shown here (many people are unmarried and some have no best friend). Where the question applies, missing data was extremely rare and virtually random (Supplementary Material APPENDIX: MISSING DATA). The scale used in subsequent analysis is the mean for each respondent on the three items, each rescored to range in equal intervals from 0 through 1

STATUS, averaged for both parents, if only one available, we use that); public worship attendance in adolescence (number of times per year at age 14), and parents' political party preference (mean of the two, each scored at equal intervals from 0 for strong Republican through 1 for strong Democrat).

The *ascriptive/ demographic traits* include age, race, ethnicity, and gender. Age is measured in single years. As a bivariate relationship, age has a curvilinear

relationship with wellbeing (e.g. Blanchflower, 2021), but in multivariate models, the functional form of the relationship is not robustly curvilinear and there is a serious danger of overfitting. In the models presented here, linearity is a good compromise. Black and Hispanic are indicator/ dichotomous variables (measured as 0=no and 1=yes). Gender is also scored as a dichotomy; female is the reference category. This variable is called "male" in the tables to signal which category is scored 1.

Current situation includes both objective indicators of opportunities and resources and several subjective measures with broad connections across different domains of life. Education is scored as years completed (e.g. Kelley & Evans, 2017), following the general principle that categorization of intrinsically quantitative variables reduces the quality of measurement, as has long been known (e.g. Bollen & Barb, 1981). Researchers do have diverse theoretical perspectives on education, but these should not detain us here as, net of the other variables in our basic model, education fails to have a significant effect on wellbeing. Occupational status (sometimes known as job quality) is scored using the VOS system ranging from 0 to 100, as detailed in Supplementary Material Appendix VERTICAL OCCUPATIONAL STATUS. Married is a dichotomous variable, scored 1 for the currently married and zero for everyone else (Supplementary Material Appendix MARRIAGE). Prior research using panel data suggests that about 2/3 of the "marriage effect" is causal, with the other 1/3 coming about via from a causal flow in the other direction (Evans & Kelley, 2004). Family income is measured in thousands of dollars, rounded to the nearest \$5,000 and transformed by a natural log. Note that "married" is included in the equation, hence controlling for that aspect of household composition. See Supplementary Material Appendix MARRIAGE for the potential effects of other aspects of household composition. Unemployment is scored as a dichotomy, with the unemployed scored 1 and all others 0. Note that income is also included in the equation, so the coefficient for unemployment will reflect the connection between the non-income facets of unemployment and subjective wellbeing. Political party preference is scored from strongly preferring the Republican party (0) to strongly preferring the Democratic party (1). Health is "In general would you say your health is... Very good, good, fair, bad, very bad".¹⁵ This was asked twice, far apart in the ~800-question surveys. The variable we use in the analysis is the average of the two instances, to reduce random measurement error (e.g., Ngamaba et al., 2017). Religious belief is a multiple-item scale (e.g., Kelley & De Graaf, 1997; Rizvi & Hossain, 2017). Time is months since the coronavirus epidemic started in mid-March.

The control variables we include do not exhaust the list of possible influences on subjective wellbeing – for example, personality (e.g., Bertrand & Mullainathan, 2001, but see recent cautions on personality measurement, Smith et al., 2021).

¹⁵ Some variants of this question in other surveys include an "Excellent" option at the top, but we have matched the EQLS formulation because, for some purposes outside of this paper, we need the same wording for comparisons.

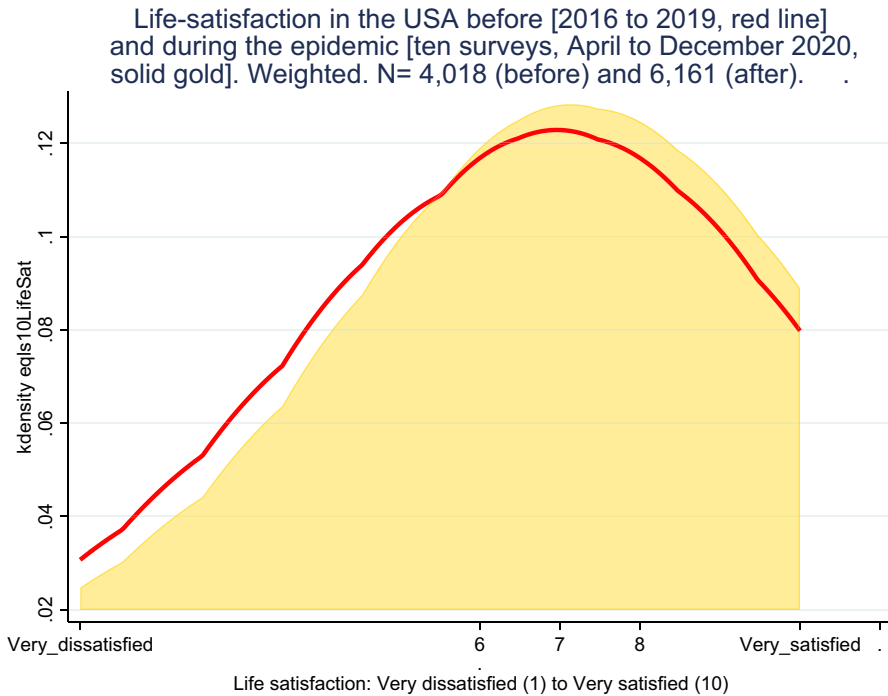


Fig. 3 Life satisfaction. Percent distributions before the epidemic (line) and during the epidemic (shape). Adult Americans, weighted $N = 4,018$ before the epidemic and $6,161$ during the epidemic

Description

Changes in Wellbeing Following the Coronavirus Epidemic

Overall, the coronavirus epidemic had (as of November 2020) had no substantial effect on Americans' *average* level of life satisfaction (Fig. 3). We have data collected in the same way (internet based) and on similar populations (two mechanical Turk samples and one from the well-regarded Survey Sampling International panel), using similar questionnaires (long, demanding, enjoyable, focused largely on income inequality) conducted a few years before the coronavirus epidemic. In these surveys (in 2016, 2017, and early 2019) we asked exactly the same questions on life satisfaction (following the wording from in the well-known European Quality of Life surveys).

The results are unambiguous: Life satisfaction early in the coronavirus epidemic (April through November 2020) was no lower than it had been before the epidemic started – indeed, if anything, it was fractionally higher (0.42 of a point on the 10 point scale; $t = 11.01$, $p < 0.001$). Moreover, the distribution was the same (familiar) shape, single humped, asymmetrical with the bulk of the population reporting moderate to high life satisfaction. In particular, there was (as yet) no sign of an unusually disadvantaged group toward the left (negative) side as might be expected from those who were

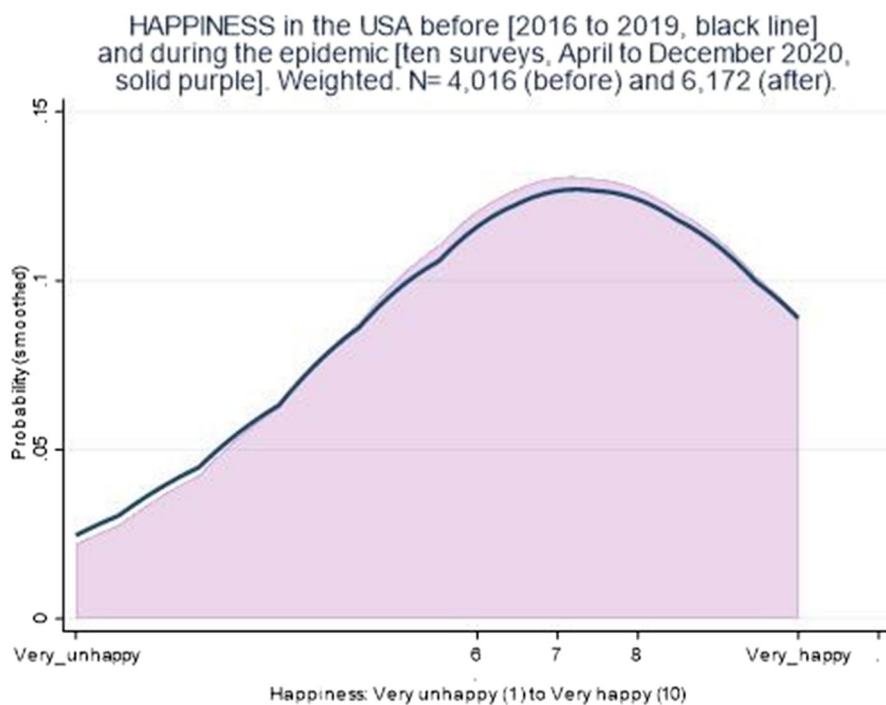


Fig. 4 Happiness. Percent distributions before the epidemic (line) and during the epidemic (shape). Adult Americans, weighted $N=4,016$ before the epidemic and $6,172$ during the epidemic

particularly anxious, or especially at risk, or particularly damaged economically – or had relatives, friends, or colleagues who were particularly unfortunate.

The same is true of happiness, also measured by a standard EQLS question (Fig. 4). In fact, the distribution of happiness before the epidemic is virtually identical to that afterwards (just 0.06 of a point higher on the 10-point scale; $t = 1.66$, n.s.).

It may be that for many people the *advantages* of the coronavirus lockdown – more time at home, less time commuting to and from work, less hassle at work, etc. – balance out (or even overshadow) the disadvantages. While this may seem surprising – given the frequent media conversation around the public tiring of restrictions – it is consistent with findings from New Zealand and South Africa, which also found no significant impact of the social distancing on wellbeing (Greyling et al., 2021a; Sibley, et al., 2020). Further research is needed to determine the reason for this somewhat surprising lack of impact on life satisfaction and identify aspects of the modifications (perhaps most likely reduced commuting) that may contribute to satisfaction, thus outweighing (on average) the costs.

Thus, the key issue is *who* has gained, who remained the same, and who lost from the epidemic. Lacking longitudinal data, we cannot directly answer the question of what individual characteristics reduced wellbeing after the lockdown

compared to before it in the present paper. Instead, we explore the correlates of life satisfaction under the current conditions, with a particular focus on the associations between pandemic-related behaviors, norms, and attitudes, on the one hand, and life satisfaction, on the other.

Social Distancing Attitudes and Behavior

Over this period, the general public's social distancing attitudes encompassed the full range from very unfavorable to very favorable (Fig. 5). A large majority of people favored distancing, many of them strongly, as shown by the tall peak well towards the favorable end of the frequency distribution. The long, thinning tail to the left of the graph nonetheless shows that dissent was by no means absent.

Similarly, large majorities of Americans reported that they were actually doing a great deal of distancing behavior (Fig. 6). Here, too, the full spectrum is represented: Most people were engaging in quite a lots of distancing behavior, but some were doing only a middling amount, yet other not doing much, and a few doing none at all.

Thus, on average Americans favor social distancing (attitude) and report that they do actually socially distance a good deal of the time (behavior). The frequency distributions of attitudes and behavior on social distancing are very similar with strong peaks towards the high/favorable end and long thin tails to the left towards the low unfavorable end. Similarly, the means are very close, with the mean on attitudes being a warm 0.83 and the mean on behavior being 82.¹⁶

Nonetheless, the correspondence between the two is far from perfect (Fig. 7). The correlation is $r=0.61$. For example, among people who strongly favor distancing (far right on the horizontal axis) some distance very little, although most do the full set of distancing behaviors regularly. Similarly, many people who moderately favor distancing (middle of the horizontal axis) also do a moderate amount of distancing, but some of them do very little and others do the full set of behaviors regularly. Among those few who are unfavorable towards distancing (far left), some do little or no distancing and others do a great deal. Thus, across the whole range of attitudes towards social distancing, we find some people doing very little distancing and others doing a great deal.

Emotional Costs of Social Distancing: Social Deprivation

Although most people approve of distancing (attitude) and in good part actually do it (behavior), that does not mean that they actually like it (evaluation). Instead for many it is bitter medicine indeed (Fig. 8).

¹⁶ To avoid clutter with leading zeros, in subsequent regression analyses independent/predictor variables (here attitudes) are scored from 0 to 1 while the dependent/response variable (here behavior) is scored 0 to 100. This is a purely stylistic difference, so just ignore the decimal points.

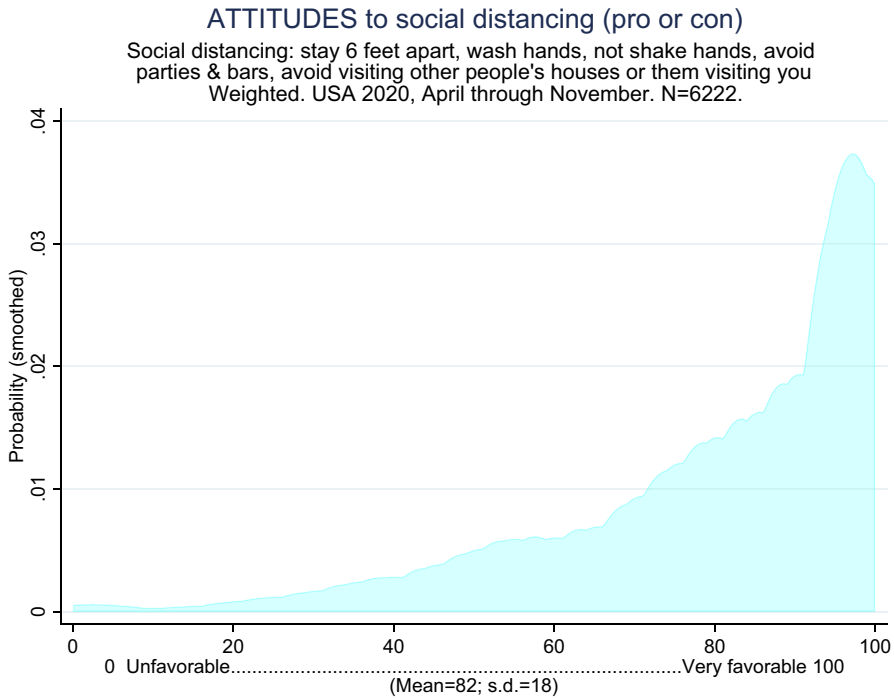


Fig. 5 Attitude towards social distancing (multiple-item scale). Smoothed percents, expressed as probability densities. Weighted data, role-players excluded. $N=6,121$ adult Americans, April–November 2020

For some people, distancing is a neutral experience or only a minor annoyance – say 0 to 20 points on the 100-point scale of emotional costliness (see those at the left side of Fig. 8, in light blue). Most people are somewhere in the middle, but for some people, distancing is horrible, in the 80–100-point range indicating great emotional cost. Note that this scale focuses on the *personal* emotional cost – how unpleasant the respondent themselves feel that taking these precautions is.

There are also real difficulties in disentangling the implications of more serious social deprivations – forgoing eating out, visits, partying, over-nighting and the like (shown toward the middle and right side of Fig. 8, in red) from more transient and usually less distressing public behaviors (staying 6 feet apart, wearing masks, not shaking hands, and the like, shown in light blue). For simplicity here we lump them all together (with a sheaf coefficient) and just estimate their combined effect. Supplementary Material Appendix EMOTIONAL COSTS takes up these matters in detail.

The social deprivation of distancing spans the full spectrum for different people, ranging from no emotional cost to great emotional cost. This is clear evidence of enormous heterogeneity in the population in the emotional costliness

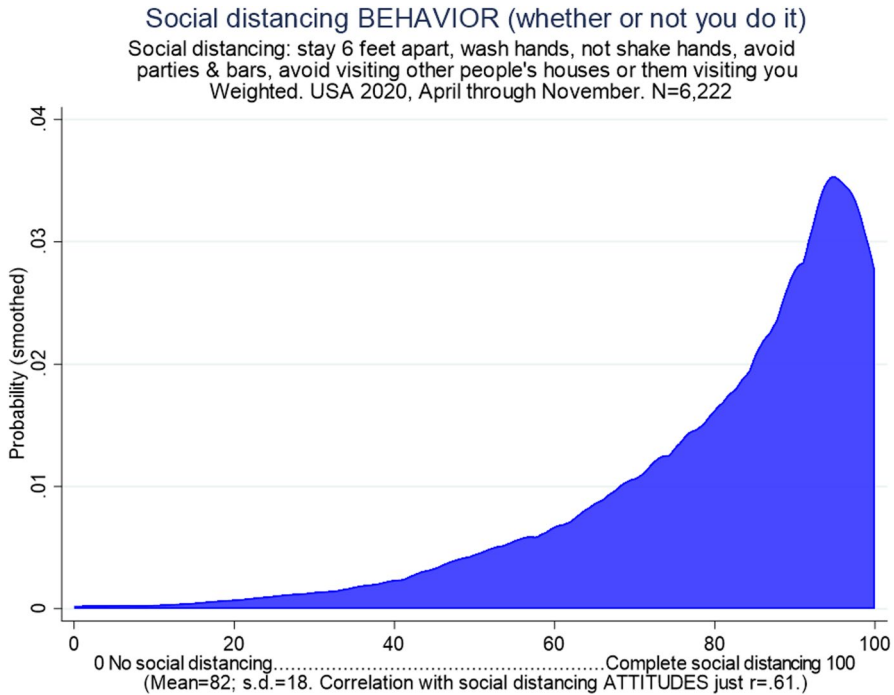


Fig. 6 Actual social distancing behavior (multiple-item scale). Smoothed percents, expressed as probability densities. Weighted data, role-players excluded. $N=6,222$ adult Americans, April–November 2020

of distancing behavior: it is bitter medicine for many, disliked by most, but no problem (or even a benefit) for a few.

Analytic Results

The results in this section are all based on ordinary least squares regression analysis with robust standard errors, weighted statistically to reflect the joint distribution of age and education (benchmark US Census Bureau) and the univariate distribution of childhood religious attendance (benchmark International Social Science Programme Religion-3). Details in Supplementary Material Appendix: WEIGHTING & ROLE-PLAYING BIAS.

The number of cases is large and missing data rare, so we confine the analysis to respondents with complete information (complete case analysis) on all relevant demographic variables and multiple-item scales. (Multiple item scales are, for each respondent, computed as the mean of all items in the scale that respondent answered. So, the scale score is missing only if a respondent failed to answer *all* items in the scale, which was exceedingly rare (Supplementary Material Appendix: MISSING DATA).

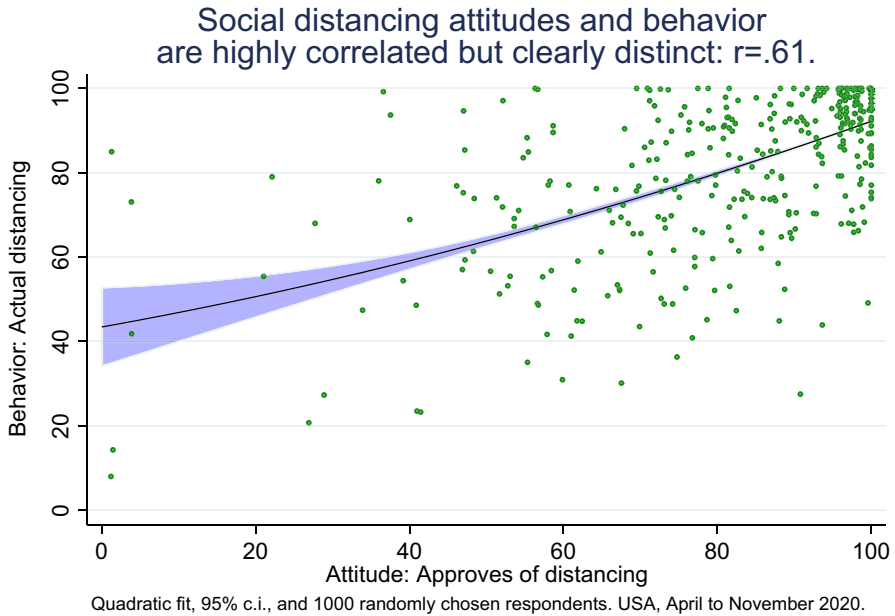


Fig. 7 Attitudes toward social distancing and actual social distancing behavior: Joint distribution (scatterplot) and quadratic fit line. $N=6,120$ adult Americans, April–November 2020. Weighted; role-players are excluded

We focus on standardized regression coefficients because they facilitate comparison across predictor variables measured in different units. e.g. years of education and dollars of income (Treiman, 2009b). They are to be interpreted as measures of the "strength" or "tightness" of a net statistical relationship. They can also be interpreted as the number of standard deviations that the response variable, on average, changes with a one standard deviation change in each predictor variable, holding the other variables in the model constant. For rules of thumb, see the "Methods" subsection of the "Methods and measurement" section above.

The samples considered here are large enough that the risk of a Type II error is trivial, so we will consider results that are not significant to be equal to zero and will focus on the relative strength of the significant results.

There are many predictor variables in our model, so it is natural to wonder if the model suffers from multicollinearity. Supplementary Material Appendix CORRELATIONS provides the bivariate correlations. Supplementary Material Appendix COLINEARITY provides the variance inflation factors (VIFs) for the regression model including all the predictor variables. VIFs under 5 are usually regarded as acceptable and even a very conservative criterion for VIFs would be that any value under 2 indicates the absence of multicollinearity. The largest VIF in our full set of independent variables is well under 2, so it is reasonable to dismiss concern about potential multicollinearity.

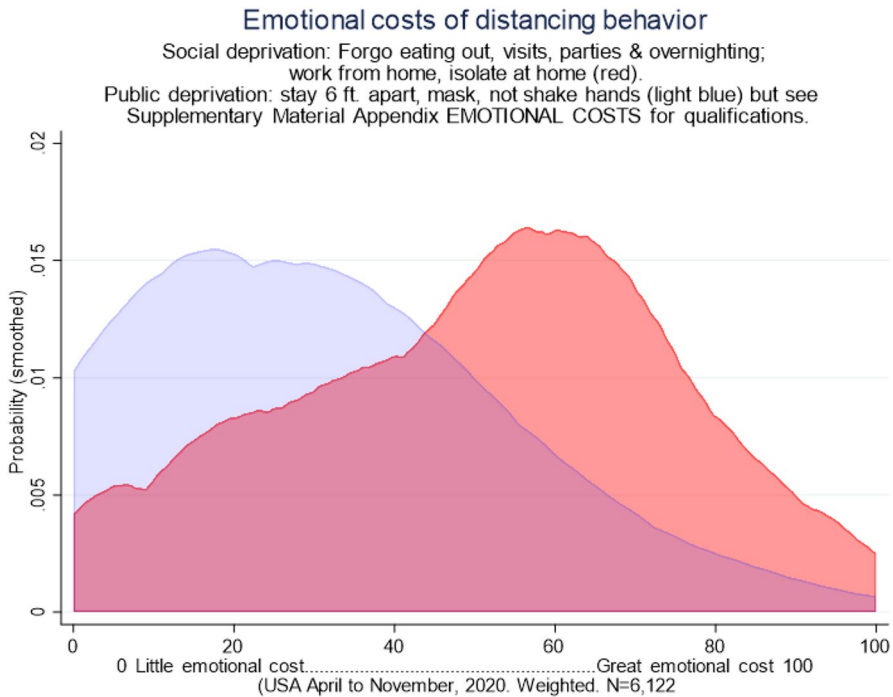


Fig. 8 Emotional costs of distancing. Smoothed percents (expressed as probability density functions). Weighted. Role-players excluded. $N=6122$ adult Americans, April–November 2020

Baseline Model

Our baseline model (Table 10, column 1, deep control variables) provides the "total effects"¹⁷ of demographic variables plus childhood background factors (e.g. Clark, 2018) from regression models with robust standard errors, as described above in the Methods section.

Most of the demographic factors have very small or non-significant net statistical relationships with subjective wellbeing: Net of other deep variables, we find that gender, race, and ethnicity have non-significant effects, men report just slightly lower levels of wellbeing than women; age has a weak but non-ignorable positive net connection to subjective wellbeing; and net of these connections, subjective wellbeing is, on average, declining very slowly over April–November 2020.

¹⁷ "Total effect" is a long-established term for the net statistical relationship of a predictor variable to the response variable in a model that does not include moderator/intermediate variables through which the influence of that predictor might flow. The term is unfortunate in light of today's strong sensitivity about causality, but it is entrenched in the literature, so we need to use it for ease of communication but to remember to interpret it tentatively. This is less of an issue with childhood and demographic predictors which are temporally long prior to our observations of subjective wellbeing.

Table 10 Influences on wellbeing (life-satisfaction, happiness): OLS regression, standardized coefficients. USA April to November 2020. Sample weighted to reflect the joint distribution of age and education (benchmark US Census) and the distribution of childhood churchgoing (benchmark ISSP Religion-3). Collinearity diagnostics in Supplementary Material Appendix COLINEARITY. Role-playing respondents and all cases with missing data are omitted (Supplementary Material Appendix WEIGHTING & ROLE-PLAYING BIAS). Total effects in red; direct effects in green; indirect effects are the difference between them

	(1)	(2)	(3)	(4)	(5)
1. Demographics					
Male (1 or 0)	0.01	-0.02	-0.01	-0.01	-0.01
Black (1 or 0)	0.03	-0.01	-0.02	-0.02	-0.02
Hispanic (1 or 0)	0.02	0	0	0	0
Age (years)[1]	0.07**	0.04*	0.04	0.03	0.04*
Time (week of epidemic)	-0.05**	-0.08***	-0.08***	-0.07***	-0.07***
2. Family background					
Parents' education (years)	0.05*	0.02	0.01	0.02	0.01
Parents' occupational status	-0.01	-0.04*	-0.04	-0.04	-0.03
Parents Democratic	-0.07***	0	0	0	0.01
Churchgoing age 14 (ln # year)	0.09***	0.03	0.03	0.03	0.02
3. Current situation					
Religious belief ($\alpha = .95$)		0.14***	0.14***	0.14***	0.14***
Education (years)		0.01	0.01	0.01	0.01
Occupational status		-0.03	-0.03	-0.03	-0.02
Married (0 or 1)[2]		0.07***	0.07***	0.07***	0.07***
Family income (ln)		0.12***	0.12***	0.12***	0.12***
Urban residence		0	-0.01	-0.01	-0.01
Subjective health (0–1)		0.36***	0.35***	0.35***	0.35***
Employed full-time		0.01	0.02	0.02	0.01
Unemployed		-0.05**	-0.05**	-0.05**	-0.05**
Democratic Party (0–1)		-0.03	-0.05*	-0.05*	-0.06**
4. Coronavirus related					
Perceived severity ($\alpha = .90$)			-0.03	-0.03	-0.04*
Attitude: Favors distancing ($\alpha = .86$)		0.11***	0.08**	0.04	
Behavior: Distances ($\alpha = .84$)				0.06**	0.04
Emotional costs: Social (stay home, no visits)			-0.08***		
Emotional costs: Public (mask, 6 ft apart) adjusted			-0.09***		
Observations	5569	5569	5569	5569	5569
R-squared	0.023	0.245	0.255	0.257	0.268

Standardized regression coefficients

[1] Age has small curvilinear effects in several columns but with a pattern that varies from column to column

[2] Other marital statuses have small effects shown in Supplementary Material Appendix MARRIAGE

* $p < 0.05$, ** $p < 0.01$, *** $p < 0.001$

Table 11 Influences on wellbeing (total effects). USA 2020. (Standardized regression coefficients, larger first)

1. Subjective health (0-1)	0.36***
2. Religious belief ($\alpha=.95$)	0.14***
3. Emotional costs of distancing (sheaf)	-0.13***
4. Family income (ln)	0.12***
5. Attitude: Favors distancing ($\alpha=.86$)	0.11***
6. Churchgoing age 14 (ln # year)	0.09***
7. Age (years)	0.07**
8. Parents Democratic	-0.07***
9. Married (0 or 1)[2]	0.07***
10. Behavior: Distances ($\alpha=.84$)	0.06**
11. Time (week of epidemic)	-0.05**
12. Parents' education (years)	0.05*
13. Unemployed	-0.05**
<u>Not statistically significant</u>	
Male (1 or 0)	0.01
Black (1 or 0)	0.03
Hispanic (1 or 0)	0.02
Parents' occupational status	-0.01
Education (years)	0.01
Occupational status	-0.03
Urban residence	0
Employed full-time	0.01
Democratic Party (0-1)	-0.03
Perceived severity ($\alpha=.90$)	-0.03

Note: Total effects from Table 6 (red). Social and public emotional costs of distancing are combined into a single variable (sheaf coefficient).

Legend

Demographics (green)

Family background (light blue)

Current situation (dark blue)

Coronavirus related (purple, highlighted)

In terms of family background/primary socialization connections, parents' education and churchgoing have weak but non-ignorable links to subjective wellbeing; parent's political party preference ranging from 0 for both strong Republicans to 1 for both strong Democrats, has a weak but non-ignorable negative net association with subjective wellbeing; and parents' occupational status does not have a significant statistical relationship to respondent's current subjective wellbeing. Even taken

together, the connections to subjective wellbeing of the predictors included in this model are weak, as indicated by an R^2 of just 2.3%.

Model 2 (in column 2) augments Model 1 with indicators of respondent's current life situation plus a few enduring subjective characteristics (current-situation control variables). Including these features lifts the R^2 to 24.5% and renders the net statistical relationship of most of the family background effects to subjective wellbeing non-significant, indicating that they are channeled through/ moderated by contemporary experiences.

Among these, subjective health stands out as having the strongest link to subjective wellbeing, *ceteris paribus*, as shown by its very strong positive standardized regression coefficient of 0.36. Moderately strong net statistical relationships connect religious belief (0.14) and family income (0.12) to subjective wellbeing. Marriage and unemployment have weak but non-ignorable net statistical connections (0.07 and -0.05, respectively). The remaining control variables representing the current situation do not have significant net statistical relationships with subjective wellbeing in these models.¹⁸

We next augment the model by beginning to add in various aspects of distancing, starting with attitudes about distancing policy (Model 3 in column 3), then adding distancing/ COVID precautionary behavior (Model 4) and finally adding the emotional impact of distancing (Model 5).

Attitudes Model 3 reveals that supporting distancing policies has a moderately strong positive net statistical relationship with subjective wellbeing, as shown by its standardized regression coefficient of 0.11.¹⁹ In order of strength, this is just behind health, religious belief, and family income in this model. In short, people who support the distancing policies promulgated by national and local health authorities and governments had higher subjective wellbeing during this period, *ceteris paribus*. Perhaps the feeling of agency and sociotropic morality— doing something collectively to slow the epidemic – engendered wellbeing. Note that this is net of political party affiliation, as well as the other control variables.

Adding in distancing behavior in Model 4, we find that practicing distancing is associated with modestly higher wellbeing, *ceteris paribus*, as shown by its standardized regression coefficient (0.06). In this augmented model, the direct connection of attitudes towards social distancing with subjective wellbeing remains present, and indeed stronger than the effect of behavior (0.08). The regression coefficients for

¹⁸ We also tested different subgroups of those not currently married (Supplementary Materials Appendix MARRIAGE). These were not statistically significant and including them did not change the regression coefficients for the focal predictors, so we did not include them in the model presented here. Note that some of the weaker effects drift in and out of significance in prior research depending on the collection of other predictors in the model.

¹⁹ We also insert perceived severity of Covid here, essentially as a control. Hypothetically it would represent a rational choice connection such that people who perceive greater danger have lower subjective wellbeing, but the connection is so weak and so non-robust, depending on what other variables are in the model that we just sideline it for purposes of this paper.

other predictor variables show little or no change with the inclusion of distancing behavior in the model.

Emotional Costs In Model 5, we turn to an aspect of the epidemic experience which has been neglected in prior research: the personal emotional cost of distancing practices. In this paper we focus on the emotional costs of two groups of activities: a public group largely concerned with the rules of interaction in public places and a social deprivation group largely concerned with forgoing social interactions (details in Supplementary Material Appendix EMOTIONAL COSTS). As shown in the Descriptive Results section above, people's feelings about these experiences span a wide range from finding them terrible to finding them mildly pleasant. The latter may seem surprising, but restrictions may also constitute a legitimated relief from social obligations to those few who find them onerous. The results in Model 5 show that both feeling that public distancing is emotionally costly and social deprivation are statistically linked to quality of life, net of all the other variables in the model. Those who find the social deprivation aspects emotionally costly have lower levels of wellbeing, as shown by the weak but non-ignorable standardized coefficient of -0.08. Many fewer people find public distancing behavior emotionally costly, and having such feeling is associated with lower subjective wellbeing (-0.09), *ceteris paribus*. Note that including these emotional costs measures in the model renders the coefficients for distancing policy attitude and distancing behavior non-significant, a result that is consistent with the interpretation that the connections between these variables and subjective wellbeing are mediated/ transmitted by emotional costs.

Summary of "Total Effects" from Baseline Models

To summarize the strength of the connections of each of our predictor variables, controlling for the other variables in the model at the stage at which the variable enters, let us consider each standardized regression coefficient before it is partialled out into indirect ("moderated") and direct linkages. For example, the "total effect" of subjective health reveals that it has the strongest connection with subjective wellbeing of any of the predictors, a standardized coefficient of 0.36 which is in the very strong range (source for this coefficient is Table 10, Column 2). The standardized regression coefficients in Table 11 are listed in decreasing order of strength (strongest first).

Religious belief comes next after health, with a moderately strong positive standardized regression coefficient of 0.14, the devout being noticeably happier everything else equal.

Strikingly, the emotional costs of distancing come next – the 3^d most important influence on wellbeing – with a moderately strong negative connection with subjective wellbeing as shown by its standardized regression coefficient of -0.12: Those who find distancing unpleasant are notably less happy, *ceteris paribus*.

The emotional costs of distancing are just as important as family income which comes next, with a positive standardized regression coefficient of 0.12. Those with

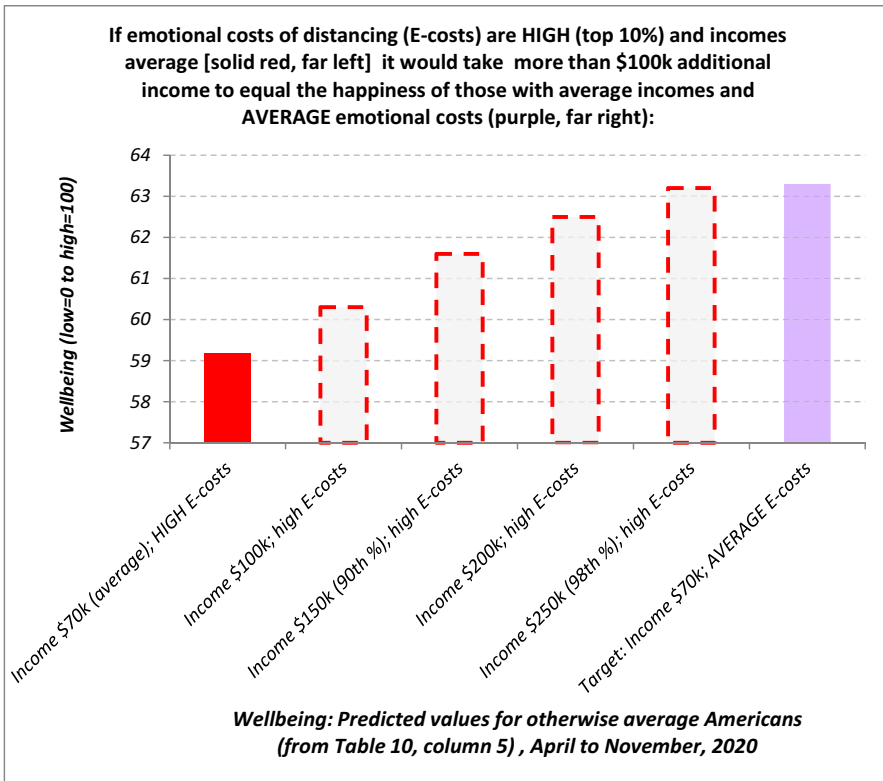


Fig. 9 Well-being by income and emotional costs

more income are of course noticeably happier. But it is striking that the benefits of income are no more important than the emotional costs of distancing.

Next, and almost as important, comes another aspect of the coronavirus pandemic: attitudes about distancing policy, with a positive standardized regression coefficient of 0.11. People who support distancing policies enjoy higher subjective wellbeing.

Thus during the pandemic, two of the top five influences on Americans' wellbeing have to do with coronavirus: emotional costs of distancing reducing it while attitudinal approval of distancing policies increases it.

Next in strength come several control variables: Churchgoing at age 14, age, parents Democrats, and being currently married.

Covid precautionary behavior – that is, actual distancing behavior – is number 10 on the list with a weak- but-non-ignorable standardized regression coefficient of 0.06. Thus, even aside from policy attitudes about distancing and aside from how they feel about distancing behavior, people who enact distancing behavior have higher subjective wellbeing, all else equal.

Of the 10 predictors with the strongest "total effects", 3 of them have to do with distancing: Feelings about distancing, attitudes on distancing policy, and

practicing distancing. This is striking in part because these results emphasize the strong linkages of transitory situations for wellbeing. Our models show a mix of connections of social and cultural characteristics with wellbeing. Some are enduring influences of socialization such as growing up in churchgoing family and the legacy of parents' political preferences. Others are the issue of the moment which everybody knows are not going to last, most notably the COVID-related subjective and objective aspects of distancing.

Commensuration: Emotional Costs of Distancing and Family Income

That having a higher family income greatly increases happiness, all else being equal, is a familiar thought and quite apparently true, as we have just seen. The same analysis showed that the emotional costs of distancing are also important, with those who find distancing unpleasant much less happy. This gives an opportunity for commensuration, putting the emotional costs of distancing into a more familiar metric of income (Fig. 9).

As a baseline, let us start with someone who finds social distancing emotionally highly unpleasant, say in the extreme 10% of the adult population on that. That comes to about 20 million Americans. According to our best estimate (from Table 10, column 5) and assuming their family income is about average (\$70,000 a year), then all else equal the best estimate of their happiness is 59 points out of 100 (Fig. 9, solid red bar on the far left).²⁰ An otherwise identical person with an only average dislike of social distancing would, we estimate, have a happiness score around 63 points out of 100 – that is to say, some 4 points higher (Fig. 9, solid purple bar on the far right).

The key commensuration question is then how much additional income it would take to move the person who finds social distancing highly unpleasant (red bar on left) to the wellbeing level of the person who dislikes it only an average amount (purple bar on the right). Bringing their income up to \$100,000 would not be enough (1st grey bar outlined in dashes); nor would \$150,000 – twice their actual income – be enough (2nd grey bar); nor even \$200,000 (3^d grey bar). Only with an income of \$250,000 would the person who finds social distancing emotionally highly unpleasant be as happy as an otherwise identical person with only \$70,000 in income but no unusual aversion to social distancing. So an additional \$180,000 (i.e. \$250,000 minus \$70,000) would compensate. There are of course substantial uncertainties in this estimate.

We conclude, conservatively, that in terms of their impact on wellbeing, some \$100,000+ a year is about the cost of feeling (as do some 20 million Americans) that social distancing is emotionally highly unpleasant. For those who have less extreme feelings about social distancing, the cost is correspondingly lower, but is still to be reckoned in tens of thousands of dollars per year.

This implies that imposing social distancing policies, however desirable they may be medically, impose heavy costs on millions of people, equivalent (in terms of reducing their happiness) to depriving them of tens of thousands of dollars each

²⁰ The standard error of this and the other estimates in Fig. 9 is around ± 2 points, so there is appreciable uncertainty in all of them.

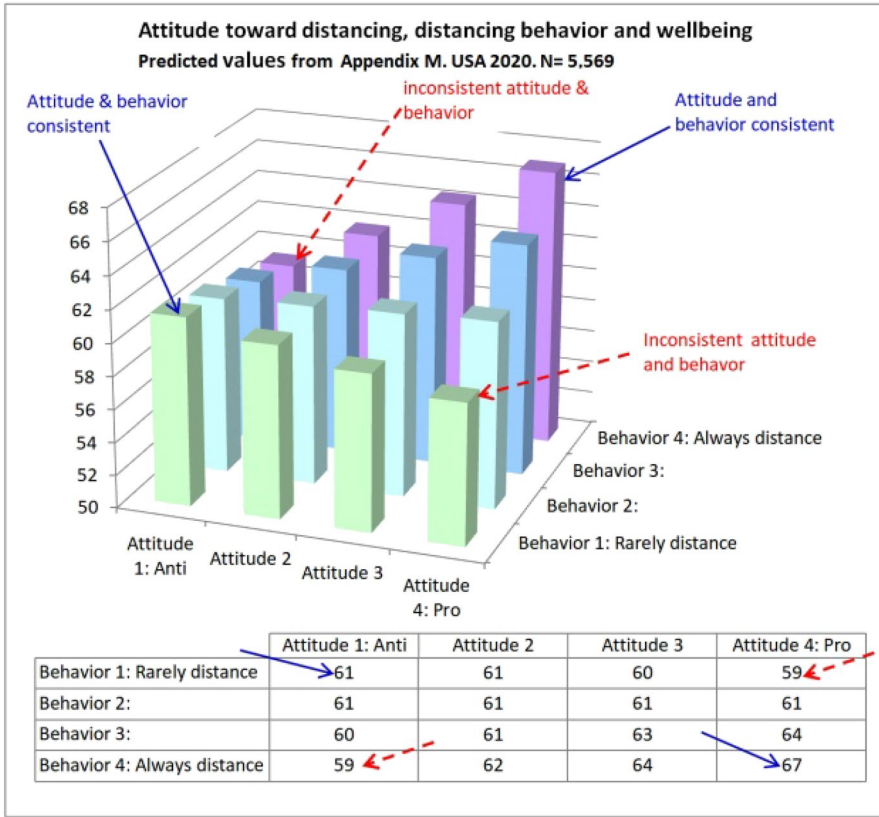


Fig. 10 Interaction effect of distancing attitude and behavior on wellbeing: Predicted values; N=5,569. Details and regression coefficients are in Supplementary Material Appendix INTERACTION

year. Unsurprisingly, many will resent it and some actively demonstrate against it. Social distancing is a bitter medicine indeed for many millions of Americans.

Final Model with Interactions: Addresses Research Question 4

The importance of the subjective and objective aspects of distancing to subjective wellbeing also raises questions of whether these are independent influences or interdependent influences, in other words, are their connections to subjective wellbeing separate or contingent on each other? As a first step in answering this question, we estimated a model that allows a multiplicative interaction (also called an intersection) between attitudes about distancing policy and one’s distancing behavior (Supplementary Material Appendix INTERACTION).

Our final model (Fig. 10) shows that there is a large and statistically significant ($p < 0.001$) interaction between *attitudes* towards social distancing and social

distancing *behavior* in their net connection to subjective wellbeing. *Ceteris paribus*, the individuals with the highest levels of wellbeing are those in the ‘aligned legitimacy’ condition: They have pro-distancing attitudes and high distancing behavior (predicted means computed from the regression model). These individuals have predicted wellbeing levels of 67 points out of 100.

All other groups have lower subjective wellbeing, all else equal. Consider, for example the individuals in the ‘coerced correctness’ condition: They do not support social distancing but are nonetheless practicing it at a high level. They have a predicted wellbeing level of 59 (fully 8 points less than their peers in the ‘aligned legitimacy’ group). Similarly, those who are ‘falling short’ (pro-distancing attitudes but low distancing behavior) could expect 59 points out of 100, all else equal. The ‘contrarians’ (anti-distancing attitudes and low distancing behavior) have a similar or fractionally higher predicted mean wellbeing level of 61.

Thus, faithful distancers may have higher or lower wellbeing than their peers depending on whether they also have a positive attitude towards distancing. This highlights the heterogeneous impact of distancing on quality of life.

Discussion

In sum, this research discovers that: (1) on average people have similar levels of wellbeing before and after the pandemic and the distribution of reported wellbeing has not changed²¹; (2) there are significant associations between attitudes about social distancing, social distancing practices, and how emotionally costly people find social distancing practices to be; and (3) the lack of change in average levels of wellbeing reflects a balance of gains (especially from favorable attitudes toward protecting people by social distancing) and losses (especially from the emotional costs of social distancing behavior).

The results shown above answer our first research question, "Do people who support distancing policies have higher subjective wellbeing, net of other influences?" with a "Yes". Despite a considerable amount of dispersion around the trend/ regression line, there is a clear positive statistical relationship between attitude towards distancing policies and subjective wellbeing: the standardized regression coefficient of 0.11 is the 5th largest effect in a model with 20 predictor variables, many of them known from prior research. People who supported distancing policies had higher subjective wellbeing during the epidemic than did otherwise similar people who had lukewarm or negative attitudes on these policies. The net connection is nearly as strong as that between family income and subjective wellbeing and slightly stronger than the well-documented association between church attendance and wellbeing.

Our results also answer our second research question “Does practicing distancing/ COVID-19 precautionary behaviors have a net positive or negative relationship with

²¹ This does not rule out different impacts of specific policies in different societies (Greyling et al., 2021b).

subjective wellbeing?" with a clear "positive". Practicing social distancing behaviors has, *ceteris paribus*, a standardized effect of 0.06 (representing higher wellbeing).

Furthermore, despite measurement challenges, the results addressing RQ3 illuminate a moderately strong connection: People who find distancing emotionally costly also have lower subjective well-being than do otherwise similar people (with a sheaf coefficient of -0.13).

A preliminary reciprocal causation model which allows feedback between subjective wellbeing and emotional costs finds an effect of emotional costs on subjective wellbeing, but no effect in the opposite direction (Supplementary Material Appendix RECIPROCAL EFFECTS). This result supports a causal interpretation that emotional costs influence subjective wellbeing, but that there is no "reverse" or "reciprocal" causation. This effect foregrounds an important aspect of social distancing that the public COVID-19 conversation is just beginning to consider: The impacts of distancing on subjective wellbeing are strongly heterogeneous, with some people finding these methods much more unpleasant than others, which impacts their overall life satisfaction.

These differences are large. Imposing social distancing policies, however desirable they may be medically, imposes heavy costs on millions of people, equivalent (in terms of reducing their happiness) to depriving each of them of tens of thousands of dollars each year. This suggests that financial incentives (or penalties) of a few thousand dollars one way or the other are unlikely to be effective. To be commensurate with the emotional costs involved requires tens of thousands of dollars per person, or even more. The only upside is that the potential gains from persuading people that social distancing is a good idea – changing their attitudes – would lead to commensurate gains in wellbeing.

Research Question 4, in turn, raise the possibility of an interaction, asking "Does the match between distancing *attitudes* and distancing *behavior* moderate the net relationship of COVID-precautionary behavior and wellbeing?". Our analysis shows that there is a strong interaction between behavior and attitudes, with those who endorse social distancing and also practice it (what we call the 'aligned legitimacy' group), net of other factors, having fully 8 points out of 100 higher level of wellbeing who do not believe in social distancing but practice it anyway (perhaps out of social or familial pressure). Those who do not believe in social distancing and also do not practice it and those who believe in social distancing but fall short of practicing it are in between these two extremes, but still have noticeably lower wellbeing than the 'aligned legitimacy' group.

This has widespread implications for both research and public policy practice. With regard to research, this illustrates the importance of understanding not only *compliance* with public health recommendations but also the *normative support* for such recommendations. On a practical level, this means that public health campaigns should work not only on convincing people to follow recommendations, but also to believe in them, as those who believe in them are not only more likely to follow them but also have substantially higher levels of wellbeing (at least in this context). As the coronavirus pandemic is far from over – and further important public health steps, such as vaccination, currently spreading slower than had been hoped– this lesson could improve and inform such campaigns.

In conclusion, these findings demonstrate that although the average level of subjective wellbeing has been surprisingly unaffected by the pandemic, people's

attitudes towards distancing, their actual distancing behavior, and the emotional cost of distancing for them all have significant effects on their quality of life. This suggests that the heterogeneous effects of public health recommendations – mediated both by perceived legitimacy and emotional cost – should be more widely considered. For instance, people are much better off following social distancing recommendations if they believe in them. Perhaps greater attention should be paid to establishing the legitimacy of new norms – as well as simply enforcing them – in order to mitigate the adverse mental health consequences of the pandemic. Forcing people to do things they do not believe in can be harmful.

Supplementary Information The online version contains supplementary material available at <https://doi.org/10.1007/s11482-023-10149-0>.

Authors' Contributions Initiative and final decisions by SK, all else equally shared.

Data Availability Data is from a proprietary survey and authors do not have permission to share publicly. Correlations, means, and standard deviations among all our basic variables are given in an appendix. Our main analysis (Table 10) and a wide variety of other analyses can readily be calculated from them.

Code Availability Stata code can be made available on request.

Declarations

Conflicts of interest No Conflicts of Interest.

References

- Adida, C., Cottiero, C., Falabella, L., Gotti, I., Ijaz, S. S., Phillips, G., & Seese, M. (2021). Taking the Cloth: How Religious Appeals Increase Compliance with COVID-19 Prevention Measures. *SocArXiv*. (Preprint server): April 20. <https://doi.org/10.31235/osf.io/yt3e>
- Andrews, F. M., & Withey, S. B. (1976). Evaluating the Measures of Well-Being. In: *Social Indicators of Well-being: Americans' Perceptions of Life Quality* (pp. 175–217). Plenum Press.
- Banks, J., & Xu, X. (2020). The mental health effects of the first two months of lockdown during the COVID-19 pandemic in the UK. *Fiscal Studies*, 41(3), 685–708.
- BBC. (2020). "Could the Swedish Lifestyle Help Fight Coronavirus? March 28, 2020" BBC, London. Retrieved Accessed Monday, August 3, 2020. <https://www.bbc.com/worklife/article/20200328-how-to-self-isolate-what-we-can-learn-from-sweden>
- Berk, R. A. (1983). An introduction to sample selection bias in sociological data. *American Sociological Review*, 48(3), 386–398.
- Bertrand, M., & Mullainathan, S. (2001). Do people mean what they say? Implications for subjective survey data. *American Economic Review*, 91(2), 67–72.
- Blanchflower, D. G. (2021). Is happiness U-shaped everywhere? Age and subjective well-being in 145 countries. *Journal of Population Economics*, 34(2), 575–624.
- Bollen, K.A. (1989). Confirmatory factor analysis. In: *Structural Equations with Latent Variables* (pp 226–318). Wiley-Interscience.
- Bollen, K. A., Biemer, P. P., Karr, A. F., Tueller, S., & Berzofsky, M. E. (2016). Are survey weights needed? A review of diagnostic tests in regression analysis. *Annual Review of Statistics and Its Application*, 3(1), 375–392.
- Brown, B. B., Clasen, D. R., & Eicher, S. A. (1986). Perceptions of peer pressure, peer conformity dispositions, and self-reported behavior among adolescents. *Developmental Psychology*, 22(4), 521.
- Bryman, A. (2016). Reliability and validity. In: *Social Research Methods* (pp. 156–162). Oxford University Press.

- Carifio, J., & Perla, R. J. (2007). Ten common misunderstandings, misconceptions, persistent myths and urban legends about Likert scales and Likert response formats and their antidotes. *Journal of Social Sciences*, 3(3), 106–116.
- Cerbara, L., Ciancimino, G., Crescimbene, M., La Longa, F., Parsi, M. R., Tintori, A., & Palomba, R. (2020). A nation-wide survey on emotional and psychological impacts of COVID-19 social distancing. *European Review for Medical and Pharmacological Sciences*, 24(12), 7155–7163. https://doi.org/10.26355/eurrev_202006_21711
- Cerqueti, R., Coppier, R., Girardi, A., & Ventura, M. (2022). The sooner the better: lives saved by the lockdown during the COVID-19 outbreak. The case of Italy. *The Econometrics Journal*, 25(1), 46–70.
- Chen, L., Savalei, V., & Rhemtulla, M. (2020). Two-stage maximum likelihood approach for item-level missing data in regression. *Behavior Research Methods*, 52(6), 2306–2323.
- Clark, A. E. (2018). Four decades of the economics of happiness: Where next? *Review of Income and Wealth*, 64(2), 245–269.
- Clifford, S., Jewell, R. M., & Waggoner, P. D. (2015). Are samples drawn from Mechanical Turk valid for research on political ideology? *Research & Politics*, 2(4), 2053168015622072.
- Cornesse, C., & Blom, A. G. (2020). Response quality in nonprobability and probability-based online panels. *Sociological Methods & Research*, 0049124120914940.
- Cornesse, C., Blom, A. G., Dutwin, D., Krosnick, J. A., De Leeuw, E. D., Legleye, S., Pasek, J., Pennay, D., Phillips, B., Sakshaug, J. W., Struminskaya, B., & Wenz, A. (2020). A review of conceptual approaches and empirical evidence on probability and nonprobability sample survey research. *Journal of Survey Statistics and Methodology*, 8(1), 4–36. <https://doi.org/10.1093/jssam/smz041>
- Daoust, J. F., Bélanger, É., Dassonneville, R., Lachapelle, E., Nadeau, R., Becher, M., ... & Stegmüller, D. (2021). A guilt-free strategy increases self-reported non-compliance with COVID-19 preventive measures: Experimental evidence from 12 countries. *PLoS One*, 16(4), e0249914.
- Evans, M. D. R., & Kelley, J. (2004). Effect of family structure on life satisfaction: Australian evidence. *Social Indicators Research*, 69(3), 303–349.
- Evans, M. D., & Kelley, J. (2008). Trends in women's labor force participation in Australia: 1984–2002. *Social Science Research*, 37(1), 287–310.
- Evans, M. D. R., Kelley, J., Kelley, S. M., & Kelley, C. G. (2019). Rising income inequality during the great recession had no impact on subjective wellbeing in Europe, 2003–2012. *Journal of Happiness Studies*, 20(1), 203–228.
- Evans, M. D. R., Kelley, J., Kelley, S. M. C. (2021). A Scale for Measuring Social Distancing Behavior: Survey Questions and National Norms, USA 2020. Working Paper.
- Gibson, J. L., Caldeira, G. A., & Spence, L. K. (2005). Why do people accept public policies they oppose? Testing legitimacy theory with a survey-based experiment. *Political Research Quarterly*, 58(2), 187–201.
- Gollwitzer, A., McLoughlin, K., Martel, C., Marshall, J., Höhs, J. M., & Bargh, J. A. (2022). Linking self-reported social distancing to real-world behavior during the COVID-19 pandemic. *Social Psychological and Personality Science*, 13(2), 656–668.
- Graham, A., Cullen, F. T., Pickett, J. T., Jonson, C. L., Haner, M., & Sloan, M. M. (2020). Faith in Trump, moral foundations, and social distancing defiance during the coronavirus pandemic. *Socius*, 6, 2378023120956815.
- Grant, N., Wardle, J., & Steptoe, A. (2009). The relationship between life satisfaction and health behavior: A cross-cultural analysis of young adults. *International Journal of Behavioral Medicine*, 16(3), 259–268.
- Greyling, T., Rossouw, S., & Adhikari, T. (2021a). The good, the bad and the ugly of lockdowns during Covid-19. *PLoS One*, 16(1), e0245546.
- Greyling, T., Rossouw, S., & Adhikari, T. (2021b). A Tale of Three Countries: What is the Relationship Between COVID-19, Lockdown and Happiness? *South African Journal of Economics*, 89(1), 25–43.
- Herby, J., Jonung, L., & Hanke, S. (2022). A literature review and meta-analysis of the effects of lockdowns on COVID-19 mortality. *Studies in Applied Economics*, 200. January 2022. <https://sites.krieger.jhu.edu/iae/files/2022/01/A-Literature-Review-and-Meta-Analysis-of-the-Effects-of-Lockdowns-on-COVID-19-Mortality.pdf>
- Holt-Lunstad, J. (2017). The potential public health relevance of social isolation and loneliness: Prevalence, epidemiology, and risk factors. *Public Policy & Aging Report*, 27(4), 127–130.
- Ikeda, T., Igarashi, A., Odani, S., Murakami, M., & Tabuchi, T. (2021). Health-Related Quality of Life during COVID-19 Pandemic: Assessing Impacts of Job Loss and Financial Support Programs in Japan. *Applied Research in Quality of Life*, 1–17. <https://doi.org/10.1007/s11482-021-09918-6>

- Jensen, U. T. (2020). Is self-reported social distancing susceptible to social desirability bias? Using the crosswise model to elicit sensitive behaviors. *Journal of Behavioral Public Administration*, 3(2). <https://doi.org/10.30636/jbpa.32.182>.
- Kauschke, C., Bahn, D., Vesker, M., & Schwarzer, G. (2019). The role of emotional valence for the processing of facial and verbal stimuli—positivity or negativity bias? *Frontiers in Psychology*, 10, 1654. <https://doi.org/10.3389/fpsyg.2019.01654>
- Kelley, J. (1990). Data, Methods, and a Worldwide Status Score. In *John Goldthorpe: Consensus and Controversy* (pp. 343–346). Falmer Press London.
- Kelley, J., & Evans, M. D. (2017). Societal inequality and individual subjective well-being: Results from 68 societies and over 200,000 individuals, 1981–2008. *Social Science Research*, 62(February), 1–23. <https://doi.org/10.1016/j.ssresearch.2016.04.020>
- Kelley, J., & Evans, M. D. (2018). Measuring occupations with VOS: A concise survey question measuring function and hierarchy. Presented to the American Association of Public Opinion Research Annual Conference. Denver, Co.
- Kelley, J., & De Graaf, N. D. (1997). National context, parental socialization, and religious belief: Results from 15 nations. *American Sociological Review*, 639–659.
- Layard, R., Clark, A. E., De Neve, J.-E., Krekel, C., Fancourt, D., Hey, N., O'Donnell, G. (2020). When to Release the Lockdown? A Wellbeing Framework for Analysing Costs and Benefits." SSRN: <https://ssrn.com/abstract=3590884> IZA Discussion Paper No. 13186(May 4):SSRN: <https://ssrn.com/abstract=3590884>
- Lawton, R. N., Gramatki, I., Watt, W., & Fujiwara, D. (2021). Does volunteering make us happier, or are happier people more likely to volunteer? Addressing the problem of reverse causality when estimating the wellbeing impacts of volunteering. *Journal of Happiness Studies*, 22(2), 599–624.
- Likert, R. (1932). A technique for the measurement of attitudes. *Archives of Psychology, Health & Medicine*, 22(1), 5–55.
- Lu, H., Nie, P., & Qian, L. (2021). Do quarantine experiences and attitudes towards COVID-19 affect the distribution of mental health in China? A quantile regression analysis. *Applied Research in Quality of Life*, 16(5), 1925–1942. <https://doi.org/10.1007/s11482-020-09851-0>
- Mazza, G. L., Enders, C. K., & Ruehman, L. S. (2015). Addressing item-level missing data: A comparison of proration and full information maximum likelihood estimation. *Multivariate Behavioral Research*, 50(5), 504–519.
- McFadden, S. M., Malik, A. A., Aguolu, O. G., Willebrand, K. S., & Omer, S. B. (2020). Perceptions of the adult US population regarding the novel coronavirus outbreak. *PLoS One*, 15(4), e0231808.
- Mills, T., & Kleinman, S. (1988). Emotions, Reflexivity, and Action: An Interactionist Analysis'. *Social Forces*, 66(4), 1009–1027.
- Motta, M., & Goren, P. (2020). Apolitical Values & Compliance with Government-Recommended Prosocial Health Behavior. *SocArXiv*. April 27. <https://doi.org/10.31235/osf.io/hxjkk5>. (Forthcoming, Journal of Elections, Public Opinion and Parties).
- Ngamaba, K. H., Panagioti, M., & Armitage, C. J. (2017). How strongly related are health status and subjective well-being? Systematic review and meta-analysis. *The European Journal of Public Health*, 27(5), 879–885.
- Nowak, P. B. (2019). Moment-type estimation from grouped samples. *SStatistics & Probability Letters*, 149, 80–85. <https://doi.org/10.1016/j.spl.2019.01.010>
- Pew. (2020). Intent to Get a COVID-19 Vaccine Rises to 60% as Confidence in Research and Development Process Increases. Pew Research Topics: Coronavirus Disease (COVID-19) 2020(Dec. 3): <https://www.pewresearch.org/science/2020/12/03/intent-to-get-a-covid-19-vaccine-rises-to-60-as-confidence-in-research-and-development-process-increases/>
- Piliavin, J. A., & Siegl, E. (2007). Health benefits of volunteering in the Wisconsin longitudinal study. *Journal of Health and Social Behavior*, 48(4), 450–464. <https://doi.org/10.1177/002214650704800408>
- Post, S. G. (2014). Altruism, happiness, and health: It's good to be good. *Positive Psychology in Behavioral Medicine*, 12(1), 66–76. https://doi.org/10.1207/s15327558ijbm1202_4
- Ravallion, M., Himelein, K., & Beegle, K. (2016). Can subjective questions on economic welfare be trusted? *Economic Development and Cultural Change*, 64(4), 697–726.
- Ridgeway, C. L. (2018). Status construction theory. In *Contemporary Social Psychological Theories* (pp. 315–339). Stanford University Press.
- Rizvi, M. A. K., & Hossain, M. Z. (2017). Relationship between religious belief and happiness: A systematic literature review. *Journal of Religion and Health*, 56(5), 1561–1582.

- Santini, Z. I., Jose, P. E., Cornwell, E. Y., Koyanagi, A., Nielsen, L., Hinrichsen, C., ... & Koushede, V. (2020). Social disconnectedness, perceived isolation, and symptoms of depression and anxiety among older Americans (NSHAP): a longitudinal mediation analysis. *The Lancet Public Health*, 5(1), e62–e70. [https://doi.org/10.1016/S2468-2667\(19\)30230-0](https://doi.org/10.1016/S2468-2667(19)30230-0)
- Sagiv, L., & Schwartz, S. H. (2000). Value priorities and subjective well-being: Direct relations and congruity effects. *European Journal of Social Psychology*, 30(2), 177–198. [https://doi.org/10.1002/\(SICI\)1099-0992\(200003/04\)30:2%3c177::AID-EJSP982%3e3.0.CO;2-Z](https://doi.org/10.1002/(SICI)1099-0992(200003/04)30:2%3c177::AID-EJSP982%3e3.0.CO;2-Z)
- Sibley, C. G., Greaves, L., Satherley, N., Wilson, M., Lee, C., Milojev, P., ... & Barlow, F. (2020). Short-term effects of the COVID-19 pandemic and a nationwide lockdown on institutional trust, attitudes to government, health and wellbeing. *American Psychological Association*, 75(5), 618–663.
- Sirkin, R. M. (2006). The standardized partial regression slope. In *Statistics for the Social Sciences* (pp. 528–530). Sage.
- Smith, M. L., Hamplová, D., Kelley, J., & Evans, M. D. R. (2021). Concise survey measures for the Big Five personality traits. *Research in Social Stratification and Mobility*, 73, 100595. <https://doi.org/10.1016/j.rssm.2021.100595>
- Springer, V. A., Martini, P. J., Lindsey, S. C., & Vezich, I. S. (2016). Practice-based considerations for using multi-stage survey design to reach special populations on Amazon's Mechanical Turk. *Survey Practice*, 9(5), 2800.
- Treiman, D.J. (2009a). Scale construction. In *Quantitative Data Analysis* (pp. 241–262). Jossey-Bass.
- Treiman, D.J. (2009b) Standardized coefficients. In *Quantitative Data Analysis* (pp. 107–110). Jossey-Bass.
- Vittinghoff, E., Glidden, D. V., Shiboski, S. C., & McCulloch, C. E. (2005). Standardized regression coefficients. In: *Regression Methods in Biostatistics: Linear, Logistic, Survival, and Repeated Measures Models* (pp. 75–76). Springer.
- Warne, R.T. (2020). "Regression." In: *Statistics for the Social Sciences* (pp. 375–410). Cambridge University Press.
- Weinberg, J. D., Freese, J., & McElhattan, D. (2014). Comparing data characteristics and results of an online factorial survey between a population-based and a crowdsourced-recruited sample. *Sociological Science*, 1(1), 292–310. <https://doi.org/10.15195/v1.a19>
- Whitt, H. P. (1986). The sheaf coefficient: A simplified and expanded approach. *Social Science Research*, 15(2), 174–189.
- Winship, C., & Radbill, L. (1994). Sampling weights and regression analysis. *Sociological Methods & Research*, 23(2), 230–257.
- Yao, R., & Wu, W. (2021). Mental Disorders Associated with COVID-19 Related Unemployment. *Applied Research in Quality of Life*, 1–22. <https://doi.org/10.1007/s11482-021-09950-6>
- Yıldırım, M., & Güler, A. (2021). Positivity explains how COVID-19 perceived risk increases death distress and reduces happiness. *Personality and Individual Differences*, 168, 110347.
- Zelditch, M. (2018). Legitimacy Theory. In P. J. Burke (Ed.), *Contemporary Social Psychological Theory* (pp. 324–352). Stanford University Press.

Publisher's Note Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.

Springer Nature or its licensor (e.g. a society or other partner) holds exclusive rights to this article under a publishing agreement with the author(s) or other rightsholder(s); author self-archiving of the accepted manuscript version of this article is solely governed by the terms of such publishing agreement and applicable law.

Authors and Affiliations

Sarah Kelley¹ · M. D. R. Evans² · Jonathan Kelley³

M. D. R. Evans
MDREvans@InternationalSurveyCenter.org

Jonathan Kelley
kelley@InternationalSurveyCenter.org

¹ Department of Data Science and Innovation, Child Trends, Bethesda, MD, USA

² Department of Sociology, University of Nevada, Reno, NV, USA

³ International Social Science Survey, Reno, NV, USA