Stigma and Health

Understanding the Psychological, Relational, Socio-Cultural, and Demographic Predictors of Loneliness Using Explainable Machine Learning

--Manuscript Draft--

Manuscript Number:	
Full Title:	Understanding the Psychological, Relational, Socio-Cultural, and Demographic Predictors of Loneliness Using Explainable Machine Learning
Abstract:	Loneliness—an important indicator of social health—is increasingly recognized to derive from factors operating at multiple levels. However, simultaneously examining the role of factors at multiple levels implies using large samples and testing multiple factors at the same time, which traditional statistical methods cannot accommodate. We used machine learning techniques to address this problem. We identify the most important out of 32 correlates of loneliness frequency in a large sample of people ages 16+ years, residing all over the world, who took part in the BBC Loneliness Experiment. Factors spanned individual, relational, socio-cultural, and demographical areas. The most statistically important associate of loneliness was daily experiences with prejudice (or stigma), followed by couple satisfaction, neuroticism (emotional stability), personal self-esteem, average hours spent alone daily, extraversion, social capital, and relational mobility. Interaction effects were also evident, showing that experiences with prejudice were most negatively associated with loneliness when individuals spent a lot of time alone, and the least when individuals were emotionally stable, had high personal self-esteem, or had high levels of couple satisfaction. This research highlights what factors need to be considered when developing effective interventions to mitigate loneliness.
Article Type:	Brief Report
Keywords:	Loneliness; machine learning; stigma; discrimination; prejudice
Corresponding Author:	Manuela Barreto, PhD University of Exeter UNITED KINGDOM
Corresponding Author E-Mail:	m.barreto@exeter.ac.uk
Corresponding Author Secondary Information:	
Corresponding Author's Institution:	University of Exeter
Other Authors:	Yiming Qin, PhD
	Christina Victor, PhD
	Pamela Qualter, PhD
Corresponding Author's Secondary Institution:	
First Author:	Yiming Qin, PhD
Order of Authors Secondary Information:	
Manuscript Region of Origin:	UNITED KINGDOM
Suggested Reviewers:	Linda Ejlskov, PhD Assistant Professor, Aarhus University: Aarhus Universitet lej@clin.au.dk Lead author of one of only two prior papers using Machine Learning to examine predictors of loneliness (restricted to older people)
	Drew Altschul, PhD Post Doctoral Fellow, The University of Edinburgh drew.altschul@ed.ac.uk Lead author of one of only two prior papers examining predictors of loneliness with machine learning techniques.

	Matthias Lasgaard, PhD Senior Researcher, Aarhus University: Aarhus Universitet au118564@ph.au.dk Leading researcher on loneliness.
	Marlies Maes, PhD Assistant Professor, Utrecht University: Universiteit Utrecht M.Maes@uu.nl Leading loneliness researcher
Opposed Reviewers:	
Order of Authors:	Yiming Qin, PhD
	Christina Victor, PhD
	Pamela Qualter, PhD
	Manuela Barreto, PhD
Manuscript Classifications:	Experienced stigma; Large data sets; Quantitative

Dear editors

We hereby submit our manuscript "Understanding the Psychological, Relational, Socio-Cultural, and Demographic Predictors of Loneliness Using Explainable Machine Learning" for your consideration in Stigma and Health. This manuscript reports on one study that uses machine learning techniques to identify the importance of 32 predictors of loneliness. We were particularly keen to highlight the role of factors that have been less well attended to, such as **stigmatization** (experiences with prejudice and discrimination). Our results indeed point to the importance of these experiences in loneliness, since they emerged as the most important of all predictors.

Although machine learning techniques are still unusual in Psychological research, we took great care to explain the technical aspects of the analyses in ways that would be understandable for the audience of this journal. This led to the use of a couple of lines more than the ideal 20 pages. We hope this is not a problem but we are very open to cutting down on words if guided by the editors and reviewers to do so. To provide complete information, but avoid unnecessary confusion, we added a file with supplementary information that can be published online along with the paper.

We very much look forward to receiving your feedback on this paper.

Best regards

Manuela Barreto
Yiming Qin
Christina Victor
Pamela Qualter

Title page with all author information

Understanding the Psychological, Relational, Socio-Cultural, and Demographic Predictors of Loneliness Using Explainable Machine Learning

Yiming Qin[a,†*], Christina Victor[b], Pamela Qualter[c] and Manuela Barreto[d]

a School of Energy and Environment, City University of Hong Kong, Hong Kong SAR, China

b Department of Health Sciences, Brunel University, London UB8 3PH, United Kingdom

c Manchester Institute of Education, University of Manchester, Manchester M13 9PL, United Kingdom

d Department of Psychology, University of Exeter, Exeter EX44QG, United Kingdom

Author note

Correspondence concerning this paper should be addressed to Manuela Barreto, Department of Psychology, University of Exeter, EX44QG, United Kingdom, email: m.barreto@exeter.ac.uk. This research was the result of a partnership between academics and All in the Mind on BBC Radio 4. We thank Claudia Hammond and Geraldine Fitzgerald for support throughout the project and Alice Eccles and Matt Richins for help with the preparation of materials. The data collection was funded by the Wellcome Trust [PQ, MB, CV, grant number 209625/Z/17/Z], who did not play any role in the study design, analysis, or interpretation of the data. The authors acknowledge that their identities can influence their approach to research. All authors identify as cisgender women, three identify as White and one as Asian, and one identifies as an immigrant. For complete research materials, data set, and data analyses scripts:

https://osf.io/9mvbk/?view_only=6497e5306e9e47bdbe270a7f82fd1d71

Understanding the Psychological, Relational, Socio-Cultural, and Demographic Predictors of Loneliness Using Explainable Machine Learning

Loneliness—an important indicator of social health—is increasingly recognized to derive from factors operating at multiple levels. However, simultaneously examining the role of factors at multiple levels implies using large samples and testing multiple factors at the same time, which traditional statistical methods cannot accommodate. We used machine learning techniques to address this problem. We identify the most important out of 32 correlates of loneliness frequency in a large sample of people ages 16+ years, residing all over the world, who took part in the BBC Loneliness Experiment. Factors spanned individual, relational, socio-cultural, and demographical areas. The most statistically important associate of loneliness was daily experiences with prejudice (or stigma), followed by couple satisfaction, neuroticism (emotional stability), personal self-esteem, average hours spent alone daily, extraversion, social capital, and relational mobility. Interaction effects were also evident, showing that experiences with prejudice were most negatively associated with loneliness when individuals spent a lot of time alone, and the least when individuals were emotionally stable, had high personal self-esteem, or had high levels of couple satisfaction. This research highlights what factors need to be considered when developing effective interventions to mitigate loneliness.

Clinical Impact Statement

This research points out the relative importance of multiple drivers of loneliness for people over 16 years old, residing all over the world. Some of the factors that emerged as most important are already often considered when developing interventions (e.g. low self-esteem), but others are less so (e.g., experiences with social stigma and poor couple satisfaction). These need to be considered by those developing interventions to prevent or address loneliness.

Key words: Loneliness, Machine learning, daily prejudice

Understanding the Psychological, Relational, Socio-Cultural, and Demographic Predictors of Loneliness Using Explainable Machine Learning

Loneliness—the feeling that one's social relationships are not as we would like them to be (Perlman & Peplau, 1981)—has significant negative consequences for individuals (Griffin et al., 2020; Park et al., 2020) and societies (Kung et al. 2021; Mihalopoulos et al., 2020), to the extent that it has been declared a public health priority in some countries (Holt-Lunstad et al., 2017) and worldwide (WHO, 2023). Despite attempts to reduce loneliness—an indicator of social health—there is little evidence of success (Eccles & Qualter, 2020; Mann et al., 2017; Quan et al., 2019). One potential reason for the modest effectiveness of existing interventions is that research seldom takes into account factors operating at multiple levels and their relative impact on loneliness. Indeed, research seldom tests the relative role of multiple predictors at the same time partly because examining the effects of a large number of predictors, and their interactions, requires large samples and complex statistical techniques. Machine Learning (ML) is one such technique and it has been recently used to examine social and psychological predictors of loneliness in samples of adults living in the United Kingdom (Altschul et al., 2019; Ejlskov et al., 2018).

Ejlskov et al. (2018) used ML to examine the combination and relative importance of 42 predictors of loneliness in a sample with 2453 participants aged 68+, from a British birth cohort study (the MRC National Survey of Health and Development). The predictors examined were personality characteristics, affective states, demographic characteristics, social relations, and health. The most important predictors of loneliness in this sample were (in this order): Positive wellbeing, personal mastery, having the spouse as the closest confidant, being extroverted, and having informal social interactions.

In turn, Altschul et al. (2019) conducted exploratory and confirmatory analyses of psychological and sociodemographic predictors of loneliness with four independent samples of older British people (45+). They examined the predictive role of personality variables, general cognitive function, subjective health and sociodemographic variables. Neuroticism and extraversion predicted loneliness among participants aged 45-69 years, and neuroticism and social circumstances (e.g., living alone) predicted loneliness among those aged 70-79.

The Current Paper

We complement existing work by using Machine Learning (ML) to identify the relative role of a series of loneliness predictors in a data set of over 40,000 individuals, ages 16-99 years, living across 237 countries, islands, and territories. These data were collected as part of a collaboration between the authors and the BBC and includes a range of variables that span multiple levels of analysis, being therefore very well suited for our goals. We extend the work by Altschul et al. (2019) and Ejlskov et al. (2018) in the following ways: (1) we include participants from a wider age range; (2) we explore a more culturally diverse sample to generalize results beyond the UK; (3) we examine a wider range of predictors that span individual, relational, socio-cultural, and demographic factors; and (4) we use an explainable ML technique that quantifies the dependencies between loneliness and the variables most related to it (i.e., interactions), while marginalizing the values of all other variables.

Regarding *individual* factors, we included both personality and wellbeing indicators. A large portion of psychological research on the predictors of loneliness has focused on individual difference factors, especially Big Five personality characteristics (Buecker et al., 2020; 2021). This research has found associations between personality variables and loneliness, particularly for neuroticism (positive) and extraversion (negative)—a finding replicated by Altschul et al. (2019) and (partially) by Ejlskov et al. (2018). Although health status is more commonly seen as an outcome of loneliness, it can also predict loneliness by decreasing a person's opportunities to engage with others (Dahlberg et al., 2021). Subjective health status was identified by Altschul et al. (2019) as one of the most important predictors of loneliness in their sample. We add to this an indicator of mental wellbeing, i.e., personal self-esteem because research has shown that personal self-esteem is a key predictor of relationship quality (Murray et al., 2002) and strongly related to loneliness (Du et al., 2018).

As to *relational* factors, both the quantity and the quality of social interactions are important determinants of loneliness (Victor et al., 2000). The quantity of social interactions (or *relational isolation*, Weiss, 1973), is often indexed by asking participants how often they meet other people, whether they live alone, or how much time they spend alone (Hawkley et al., 2005). In addition, individual's attitudes towards living alone (including whether this is a choice), and even their evaluation of loneliness experiences, can predict how prevalent loneliness is in their lives (Wang et al., 2013).

Although indicators of the quantity of social interactions are relatively straightforward, and often included in research, indicators of the quality of a person's social interactions are often left out (possibly in part due to concerns about their overlap with measures of loneliness) or limited to interpersonal relationship quality. For example, Altschul et al. (2019) did not include relationship quality in their predictors while Ejlskov et al. (2018) asked participants to indicate the level of emotional support they received from the person they felt closest too, and negative aspects of this relationship. Although this seems important, loneliness can also be predicted by the quality of daily interaction experiences with others with whom one does not necessarily have a close relationship (Cacioppo & Cacioppo, 2012). For example, there is evidence that daily interpersonal experiences with prejudice and discrimination (or social stigma) are important determinants of loneliness (Lee & Bierman, 2019; Priest et al., 2017), and that positive and trusting relationships with one's neighbors (corresponding to high social capital) can protect against loneliness (Matthews et al., 2019). Therefore, we indexed relationship quality through couple satisfaction, daily experiences with prejudice, and neighborhood social capital.

Socio-cultural differences in individualism-collectivism (Hostede, 1991; Triandis, 1995)—reflecting the extent to which a given society values loose vs. tightly knit networks—can also impact loneliness, although evidence is mixed regarding the direction of this effect. Another cultural variable that might be relevant in this context but has only been examined among adolescents (Jefferson et al., 2023a), is power distance (Hofstede, 1991), consisting of the extent to which a social environment promotes the existence of power differences between people, or whether it strives for more egalitarian relationships. Finally, researchers have examined the impact of relational mobility—the extent to which social relationships in each network or society tend to be primarily chosen or ascribed (Yuki & Schug, 2020)—on various aspects of social networks, but the impact of this variable on loneliness is yet to be examined.

Research has also shown that certain *demographic* characteristics are associated with loneliness. This research has examined effects of age, gender, educational level, and socio-economic or employment status (Buecker et al., 2020a). Less frequently, researchers have demonstrated that some demographic characteristics associated with social roles (such as being a carer, or a parent of young children) or with socially stigmatized characteristics (homelessness, minority sexual orientation, migrant status) can make people vulnerable to loneliness. Contrary to what is commonly assumed, loneliness is not most prevalent in older people, with studies that include samples with a wide age range showing that young people 16-25 report the highest levels of loneliness (Barreto et al., 2020; ONS, 2018). Effects of gender are inconsistent, with a meta-analysis showing that, overall, these are small and generally negligible (Maes et al., 2019). In addition, socially stigmatized groups experience more loneliness than non-stigmatized groups (see also Barreto et al., 2023). For example, both young (Madsen et al., 2016) and older (Victor et al., 2012) migrants report more loneliness than those without a migration experience; individuals with a mental illness report more loneliness than those without a mental illness (Lauder et al., 2004); and sexual minority individuals report more loneliness than heterosexuals (Doyle & Molix, 2016). High levels of loneliness have also been reported by individuals with low socioeconomic status (Morgan et al., 2019), homeless youth (Kidd, 2007), individuals with a disability (Tough et al., 2017), and unemployed individuals (Kleftaras & Vasilou, 2016). To account for such possible loneliness discrepancies, we examined the role of a range of demographic characteristics.

Studies focusing on a small number of predictors at the time are important, but they do not allow for the simultaneously examination of factors operating at multiple levels, and their interactions, to shed light on the relative importance of each factor. They also involve substantial subjectivity in deciding what variables and interaction terms to include in the analyses, and in what order, as well as the risk of multicollinearity. However, advanced machine learning techniques, which capture patterns from data, can handle those challenges a lot better. While there will always be some degree of subjectivity involved in the selection of variables on which to collect data, this technique identifies the interactions that are useful to examine from the data itself. Given lack of consensus in the prior studies that have used this method, the wider age and cultural diversity in our sample, and the different indicators, we did not raise specific hypotheses for this study about the relative importance of predictors of loneliness in this study, which remained exploratory.

Method

We used cross-sectional data from the BBC Loneliness Experiment. Data were collected online in 2018, from participants aged 16-99 years living in one of 237 countries, islands, and territories (Barreto et al., 2021). The sample was recruited over a month period without aiming for a predetermined sample size. We use data from all participants who had data in the measures of interest, resulting in a sample size of 40,080. The characteristics of the sample can be seen in Table 1.

Loneliness was measured with four items from the UCLA Loneliness Scale (Russell, 1996): do you feel a lack of companionship?, do you feel left out?, do you feel isolated from others? and do you feel in tune with people around you? (reverse coded). Each item was rated on the frequency with which it was true for the participant (from 0 = never to 5 = always; internal reliability: Cronbach's $\alpha = .84$).

We measured personality, using the 10-item scale by Gosling et al. (2003), which has adequate internal reliability ranging from .48 for 'Openness to Experience,' to .71 for 'Emotional Stability' (Burns et al., 2017). Wellbeing was measured with one indicator of psychological wellbeing (personal self-esteem, measured with four items from the Rosenberg, 1965, scale, e.g., "On the whole, I am satisfied with myself", $\alpha = .91$) and one item measuring subjective health ("Would you say that, in general, your health is", from 1 = poor to 5 = excellent).

Quantity of social contact was indexed in several ways. First, participants indicated if they lived alone, and if so how long (in months); participants who said they did not live alone were asked: "How many people (excluding yourself) live in your household?" (open answer); all participants were asked "How much time do you spend alone?" (from 1 = never to 4 = always), and "on average, how many hours do you spend alone in one day?". To assess participants' attitudes towards living alone and towards loneliness, we asked: "Did you choose to live alone?"; "how much do you enjoy spending time alone?"; and "Is the experience of loneliness positive for you?" (no, sometimes, yes). The latter question was not shown to participants who indicated never feeling lonely.

Regarding quality of social contact, couple satisfaction was measured with the four-item version of the Couples Satisfaction Index (Funk & Rogge, 2007). This measure was only presented to participants who indicated being in a relationship. An example item is: “How rewarding is your relationship with your partner?”, from 1 = not at all to 7 = completely; α =.94). Participants’ daily experiences with prejudice and discrimination were assessed with the five-item version of the Everyday Discrimination Scale by Sternthal et al. (2011). Participants were asked to indicate how often each of the five items happened to them. A sample item is: “You are treated with less courtesy or respect than other people” (from 1 = never to 7 every day; α =.79). Social capital was measured with the seven-item scale by Martin et al. (2004), with items such as “People around my local neighborhood are willing to help their neighbors” from 1 = strongly disagree to 5 = strongly agree; α =.82).

To operationalize collectivism and power distance, participants indicated their country of residence, which was coded using Hofstede’s (1997) indices, with 0 corresponding to collectivism or low power distance and 100 corresponding to individualism and high power distance. Our participants resided in countries that spanned the full range of these two dimensions. We also measured relational mobility with the 12-item scale by Thomson et al. (2018). Participants were asked to reflect about the people in their immediate society and to indicate to what extent they agreed with each item. An example item is “They are able to choose, according to their own preferences, the people whom they interact with in their daily life” (1 = strongly disagree to 6 = strongly agree; α =.90).

Demographic information provided gender (male, female, other, prefer not to say); age; marital status (single, in a relationship but not living together, married or cohabiting, separated or divorced, widowed); country of residence; country of birth; employment status (retired, in part or full time work, part or full time student, unemployed); education level (years of education completed); income ("how well do you feel that your needs are met by the financial resources you have?", very well, fairly well, poorly); subjective socio-economic status (MacArthur scale, Adler et al., 1994; from 1 = bottom rung and 10 = top rung); carer (yes/no); if participants indicated being a carer, they also indicated how long they had been a carer; dependents (yes/no); number of children and age of youngest child (both only for those with children); and sexual orientation (from 1 = exclusively heterosexual to 6 exclusively homosexual, with 7 = asexual). Migrant status was computed using birthplace and place of residence, by categorizing participants as living in the country of their birth (1) or not (0).

Ethical approval was obtained for this study prior to data collection from the University Research Ethics Committee at the University of Manchester.

Analytical Strategy

The analyses reported in this paper were not pre-registered. We used Machine Learning (ML), which involves the searching for generalizable patterns to make precise predictions from a dataset. ML contrasts with traditional statistics that focus on inferring relationships between variables from a sample. ML models provide four advantages compared to traditional statistical methods (Kyriazos et al., 2022): (1) no assumptions about the distribution of the dependent and independent variables need to be made, (2) ML uses training data to recognize patterns and make predictions to be tested in test data, (3) it manages missing data effectively, and (4) it can handle large datasets efficiently. To identify the most important factors related to loneliness, we used Random Forest analysis, which is based on the results of an ensemble of regression trees to predict the response values. Random Forests can effectively model complex non-linear relationships between input features (i.e., predictors) and the target variable through the collection of decision trees. Each tree makes decisions based on thresholds in features splitting the input space into piecewise-constant segments. For example, if the decision tree determines that ‘6 hours alone per day’ is an important point at which loneliness changes, then this becomes a threshold. This piecewise approximation allows Random Forests to adapt to data with multiple interaction effects and high-dimensional feature spaces without the need for explicit feature transformation, making them powerful for capturing non-linearity in regression tasks. In this study, we use Random Forest to analyze the relationship between loneliness frequency and the other variables.

The standard practice of allocating 80% of the data for the training set and the remaining 20% for the test set (Joseph, 2022) was adopted. During the model training process, hyperparameters were selected by minimizing the mean squared error (MSE). Predictions on the test set were made using these optimal hyperparameters. The feature of importance was computed by averaging the reduction in mean squared error attributed to each feature across all trees, representing the relative importance or contribution of each feature to the prediction model.

In a second stage of the analyses, partial dependence plots (PDPs) were employed to assess how the most important predictors influence the prediction of loneliness when all other variables in the model are held fixed. The partial dependence plot algorithm was proposed by Friedman (2001) for investigating the relationships among input variables and the output prediction. The advantage of the partial dependence plot compared to that of the conventional regression on the scatter plots is that partial dependence plots allow us to visualize how relatively small changes in the predictor are associated with changes in the outcome variable, while at the same time excluding the effects of other confounding predictors through a marginalized distribution (a detailed explanation of the partial dependence plot can be found in Qin et al., 2022). This method allows for the individual effects of each predictor on the outcome (loneliness) to be examined in detail and in isolation of other predictors.

We employed both 1-D and 2-D PDPs. A 1-D Partial Dependence Plot (1-D PDP) visualizes the relationship between a single feature and the predicted outcome (loneliness frequency), by plotting the prediction against different values of the feature while other features are held constant. For the 1-D PDPs we selected as the x variable only the predictors that explained at least 5% of variance in loneliness frequency and all other variables were held at their respective mean values. In addition, we employed 2-D PDPs to show the interaction between two features and how they jointly influence the prediction of loneliness frequency. For the 2-D PDPs we included as predictors the variable that emerged as the most important predictor in interaction with the remaining post important predictors (those who explained at least 5% of variance). The complete research materials, data set, and data analyses scripts can be found here: https://osf.io/9mvbk/?view_only=6497e5306e9e47bdbe270a7f82fd1d71

Results

Loneliness frequency was widely spread across the scale (from 1 to 5), with a mean of 2.66 and a standard deviation of 1.13 (see Table 1 and Figure S1). The correlation coefficients (R^2) prediction for loneliness frequency in the training set and the test set were 0.93 and 0.48. respectively. The Random Forest model exhibited a high degree of accuracy on the training set, accounting for 93% of the variation in loneliness frequency, which suggests a strong alignment with the training data. On the test set, the model explained 48% of the variation, indicating a moderate predictive performance on unseen data. It is worth noting that it is normal for the training set to have a much higher degree of accuracy because the model is specifically tuned to this data, whereas the test set is tested in unseen data (James et al., 2013). This suggests that while there is room for improvement in the model performance, it has a considerable amount of predictive power when applied beyond the data it was trained on. Both values were above the value reported by Ejlskov et al. (2018), which was 32% accuracy, suggesting that the variables included in this study add predictive power to those examined in prior research.

Figure 1 shows the important identifiers of loneliness frequency in Random Forests. The variable importance measure was scaled so that the sum of all feature importance scores becomes 100%, providing the relative importance of a variable among all input variables. Higher values of importance indicate that those features have a higher impact on loneliness. The wide spread of the important identifiers indicates that loneliness was associated with several variables. Among the 32 variables included in the analyses, daily experiences with prejudice and discrimination, couple satisfaction, emotional stability (also called neuroticism), self-esteem, hours spent alone daily, extraversion, social capital, and relational mobility were identified as the most important variables, each accounting for more than 5% of variance in loneliness. This means that each one of these variables individually contributed substantially to the differences observed in loneliness levels across the study sample, making them particularly important for understanding or predicting loneliness. As can be seen in Figure 1, the next strongest predictors explained 3% or less variance in loneliness.

To further explore the relationship between loneliness frequency and these eight most important predictors, we estimated the partial dependence of loneliness frequency with respect to *changes* in experiences with these predictors using 1-D Partial Dependence Plots (see Figure 2). The changes in the specific input variable (i.e., each of the eight predictors that explained at least 5% of variance in loneliness frequency) in relation to loneliness frequency was estimated while considering the variability of the rest of the variables through marginal effects. In this way, we could inspect the expected loneliness frequency as a function of the input features of interest. The partial dependence of loneliness frequency with experiences with prejudice indicates that there was a relationship between the frequency of loneliness and these experiences, but this relationship was not constant across all levels of prejudice. When experiences with prejudice were at the lowest measured level, represented as level 1, and then increased to level 2, the frequency of loneliness experiences only slightly increased, suggesting a relatively stable or weak response to low levels of prejudice. However, as the level of experienced prejudice rose to a moderate level (level 2) and continued to increase to a higher level (level 4), there was a pronounced surge in the frequency of loneliness. This implies that the impact of discrimination on loneliness became significantly stronger as experiences with prejudice crossed a certain threshold. Beyond that point, increases in experiences with prejudice were associated with more substantial increases in feelings of loneliness.

In terms of the relationships between loneliness frequency and couple satisfaction, the results indicate a clear link between couple satisfaction and the frequency of loneliness. Specifically, as couple satisfaction scores improved, up to a score of 20, there was a noticeable decline in the frequency of loneliness, dropping from an average of 3 to 2.5. Once couple satisfaction exceeded a score of 20, the frequency of loneliness stabilized and did not decrease further, suggesting that beyond this point, increases in couple satisfaction did not have a significant impact on reducing loneliness frequency.

The relationship between loneliness frequency and emotional stability was characterized by a negative correlation, where an increase in emotional stability was associated with a reduction in loneliness frequency. This reduction in loneliness was more pronounced when emotional stability levels rose from a lower score of 2 to a higher score of 5, indicating that individuals with greater emotional stability were likely to experience loneliness less frequently, especially as they moved from lower to higher levels of emotional stability.

In terms of self-esteem, as self-esteem increased, particularly within the range of 15 to 20, there was a corresponding decrease in the frequency of loneliness. This suggests that improvements in self-esteem were associated with a decline in how often individuals felt lonely. We also observed an extreme value when self-esteem is at 0, indicating a significant spike in loneliness frequency. This means that at the lowest possible level of self-esteem, the frequency of loneliness was notably higher compared to at other levels of self-esteem, underscoring a strong inverse relationship between self-esteem and loneliness, where very low self-esteem was associated with much higher occurrences of loneliness.

We also observed the frequency of loneliness experiences a slight rise as the number of hours a day spent alone (on average) increased. This increase became more pronounced when the time spent alone extended beyond 7 hours a day, particularly when it surpassed 20 hours. Therefore, while there was a general trend for more loneliness with increased alone time, it is especially noticeable at higher thresholds of time spent alone.

The 1-D PDP plot also shows that loneliness frequency gradually decreased when extraversion increased. This suggests that more extroverted individuals generally reported feeling lonely less often than their less extroverted counterparts. Similar to the effect of extraversion, an increase in social capital also corresponded to a decrease in the frequency of loneliness. This suggests that individuals who had more social resources at their disposal were less likely to experience loneliness, emphasizing the importance of social connections and community involvement in mitigating feelings of isolation. Likewise, with relational mobility, as relational mobility increased, the frequency of loneliness tended to decrease. This indicates that in environments where individuals have more opportunities and feel more at ease to establish and change social connections, they were generally less likely to report feelings of loneliness. This underscores the value of being in a dynamic social environment that supports and encourages the formation of new social ties.

Given that experiences with prejudice and discrimination were identified as the most critical factor for predicting loneliness, as illustrated in Figure 1, we further examined how prejudice interacts with other factors to predict loneliness. To investigate this, we generated 2-D Partial Dependence Plots. A 2-D partial dependence plot depicts in detail the interaction between (changes in) two predictors and their combined effect on a response variable, with other variables held constant at their average values. It displays this interaction on a grid where the axes represent the predictors and the surface color indicates the predicted outcome (here, loneliness frequency). The plot uses colors to represent different levels of predicted loneliness across two dimensions: One axis for prejudice and the other for the second variable of interest. By interpreting the colors, we can discern how changes in prejudice levels and this second variable together influence the frequency of loneliness, with the color intensity typically indicating higher or lower values of the predicted outcome. The lighter color in Figure 3 indicates a higher value of loneliness frequency.

Figure 3 suggest that the impact of prejudice on loneliness was influenced by several other factors, including emotional stability, couple satisfaction, hours spent alone, and personal self-esteem. First, emotional stability appears to buffer the effects of prejudice; people who had higher emotional stability experienced a lower frequency of loneliness even when they faced similar levels of prejudice compared to those with lower emotional stability. Second, the relationship between couple satisfaction and loneliness in the context of prejudice is more complex. When couple satisfaction is low to medium, it doesn't significantly alter the loneliness frequency associated with prejudice. However, at higher levels of couple satisfaction, there is a notable decrease in loneliness frequency, even with the same experiences of prejudice. This indicates that high couple satisfaction can mitigate the negative effects of prejudice on loneliness. Third, hours spent alone interacted strongly with prejudice experiences. Regardless of the level of prejudice faced, the more time participants spent alone, the higher their frequency of loneliness. This suggests that spending time alone can amplify the loneliness that prejudice can cause. Lastly, personal self-esteem also played a protective role. Individuals with higher self-esteem experienced less loneliness at equal levels of prejudice experiences than those with lower self-esteem. This suggests that higher self-esteem can reduce the negative impact of prejudice on loneliness frequency. These interactions imply that emotional stability, couple satisfaction, time spent alone, and self-esteem were significant moderators of the relationship between prejudice experiences and loneliness, either weakening or strengthening the effect of prejudice experiences on how frequently individuals feel lonely.

Exploratory analyses (described in the Supplementary Materials) additionally show that the association of loneliness with prejudice, couple satisfaction, emotional stability and extroversion are similar for all age groups.

Discussion

We used advanced ML to explore the relative importance of 32 individual, relational, socio-cultural, and demographic factors as correlates of loneliness frequency among those who participated in the BBC Loneliness Experiment. While examination of the correlates of loneliness is not new, the identification of the unique and most important associations of loneliness among those aged over 16 years, using ML, is novel. By expanding the age range of previous ML samples (Altschul et al., 2019; Ejlskov et al., 2018) to adolescence (from 16 years) and young adulthood, we were able to explore important correlates of loneliness across ontogeny. We also examined a more culturally diverse sample than in previous ML work, and focused on individual factors, relational, socio-cultural, and demographic factors. Our findings support existing ML research on loneliness and earlier work in the field that used traditional statistical methods, showing there to be important associates of loneliness, but our Random

Forest ML provides more precision compared to those earlier studies that used regression analyses and advanced the previous work using ML.

We found that the main correlates of loneliness among those aged 16 and 99 years were, in order of importance: Everyday experiences with prejudice, couple satisfaction, emotional stability (neuroticism), average hours spent alone per day, low self-esteem, extraversion, social capital, and relational mobility. This highlights the need to examine multiple factors simultaneously and provides information that can be used to inform the development of interventions, thus, expanding both our understanding of loneliness and the potential avenues for support.

The most important associate of loneliness for people who participated in the BBC Loneliness Experiment was daily experiences with prejudice and discrimination, suggesting that loneliness often results from processes of social marginalization and devaluation (see also Barreto et al., 2023 and Barreto et al., in press). Importantly, it was not the demographic membership of groups that are often marginalized (such as migrants) that predicted loneliness, but daily experiences with prejudice. This clarifies that loneliness often emerges from processes of exclusion, rather than being inherent to how members of these groups function socially. These findings suggests that the most needed interventions to reduce loneliness are those focused on making social environments more inclusive (see also Jefferson et al., 2023a,b). Enhancing our understanding of the ways in which prejudice and discrimination can affect loneliness requires greater research attention to understand whether specific types of discrimination are related to loneliness and how that association works prospectively.

Across several studies using traditional statistical methods, couple satisfaction was a powerful factor protecting individuals from reporting loneliness (Luhmann & Hawkley 2016). We support that finding using more sophisticated ML. Recent prospective work has shown that loneliness can also predict couple satisfaction (Mund & Johnson, 2021). Our finding that this variable emerged as one of the most important associates of loneliness further underlines the relevance of further exploring this bidirectional relationship.

A recent meta-analytic study (Buecker et al., 2021) showed that emotional stability was the strongest predictor of loneliness among the big five personality traits. Our findings align well with that work, but the mechanisms linking emotional stability and loneliness still need to be examined. There is evidence that this variable is related to a heightened reactivity to social stressors (Zautra et al., 2005), but also that individuals low in emotional stability are more sensitive to social rejection cues (Denissen & Penke, 2008), both of which are linked directly to increases in loneliness (Qualter et al., 2015). Direct examination of the mechanisms linking neuroticism and loneliness is needed. Similarly, we found that extroversion was one of the most important associates of loneliness and that its effect was similar for all age groups and for men and women.

While aloneness is not the same as loneliness, we found the number of hours a person spent alone on an average day were an important associate of loneliness, when all others were controlled for, particularly when people were spending quite a lot of time alone (more than 20 hours). Indeed, spending time alone can be very valuable, and often deliberately sought, but spending a lot of time alone on a daily basis can be detrimental to wellbeing. This finding lends complexity to the commonly expressed idea that aloneness and loneliness are distinct and suggests that more research is needed to better understand the relationship between these variables, how much time alone is detrimental, and how people might be able to monitor their alone time to avoid loneliness.

Low self-esteem also emerged as an important correlate of loneliness, when all other variables were kept constant, which is consistent with a wide range of studies linking self-esteem to relational behavior and relationship quality (e.g., Murray et al., 2002). Low self-esteem can perpetuate loneliness by increasing hypervigilance, biasing interpretations of others' behaviours as rejecting, leading to defensive behaviours and motivating withdrawal (Qualter et al., 2015). Importantly, self-esteem can be lowered by a variety of life experiences, such as bullying and other forms of victimization, which points towards areas of intervention to reduce experiences that lower self-esteem.

Importantly, our analytical technique and sample size allowed us to examine interactions between variables. While computing interactions between all possible variables would be unfeasible, we examined interactions between the variable that emerged as most important (experiences with prejudice) and the remaining variables that explained at last 5% of variance. The results indicate that experiences with prejudice have the most negative effects on loneliness when individuals spend a lot of time alone, and the least when individuals are emotionally stable, have high personal self-esteem, or have high levels of couple satisfaction. It seems reasonable that if people have few social experiences, those they have need to be positive, justifying why experiences with prejudice are particularly problematic when people spent a lot of time alone. How self-esteem protects from the emotional of prejudice and discrimination had been established in a series of experimental studies, but this had not yet been done in connection to loneliness (Cihangir et al., 2010). In turn, while couple satisfaction has been shown to be detrimental affected by experiences with prejudice (e.g., Doyle & Molix, 2014), its role as a moderator of the impact of these experiences on loneliness had not yet been demonstrated. The role of emotional stability as a moderator of the impact of prejudice has, to our knowledge, not been documented before. Importantly, experiences with prejudice had a similar association with loneliness frequency for all age groups and for men and women. Supplementary analyses also showed that couple satisfaction, emotional stability, and extroversion were similarly associated with loneliness at all ages and for men and women.

It is important to acknowledge that even though our machine learning approach provides high levels of precision and nuance, these findings are ultimately based on correlational data and that further research is needed to examine these associations prospectively and experimentally to clarify causal relationships. That said, the analytical techniques we employed do go further than prior research by identifying particularly important associations, when multiple others are controlled for, as well as enabling the more detailed examination of how change in a predictor is associated with changes in loneliness. In this way, our findings point to several factors that are already targets for intervention, many having been shown to be moderately successful at reducing loneliness (Lasgaard et al., 2022), but adds by highlighting which of these require most attention, as well as by stressing some factors that are not yet receiving enough attention in loneliness interventions (e.g., stigmatizing experiences). In addition, the robustness of the ML and the inclusion of a wider range of ages and cultures than in prior research offer precision and superior evidence; the inclusion of a measure of everyday discrimination (which was not included in prior research using ML), and the fact that it was the strongest predictor of loneliness by far, offers new ideas for intervention that would work well alongside more traditional intervention strategies focused on relational and individual changes.

Conclusions

Our findings show that the key correlates of loneliness, when others are kept constant, are socio-cultural (discrimination), relational (couple satisfaction, hours spent alone), and individual (neuroticism, personal self-esteem). As such, interventions need to focus on multiple factors and, crucially, address marginalization. Indeed, the typical focus on individual and relational strategies without addressing structural factors will do little to mitigate loneliness and the adverse effects that it has on health and well-being, creating further inequalities for already marginalized groups.

References

Adler, N. E., Boyce, T., Chesney, M. A., Cohen, S., Folkman, S., Kahn, R. L., & Syme, S. L. (1994). Socioeconomic status and health: The challenge of the gradient. *American Psychologist*, *49*(1), 15-24. 10.1037//0003-066x.49.1.15

Altschul, D., Iveson, M., & Deary, I.J. (2019). Generational differences in loneliness and its psychological and sociodemographic predictors: An exploratory and confirmatory machine learning study. *Psychological Medicine*, 1-10. https://doi.org/10.1017/S0033291719003933

Barreto, M., Doyle, D., & Qualter, P. (2024). Changing the narrative: Loneliness as a social justice issue. *Advances in Political Psychology*. http://doi.org/10.1111/pops.12965

Barreto, M., Doyle, D.M., & Qualter, P. (2023). *Social inequalities in loneliness*. Welsh Centre for Public Policy.

Barreto, M., Victor, C., Hammond, C., Eccles, A., Richins, M., & Qualter, P. (2020). Loneliness around the world: Age, gender, and cultural differences in loneliness. *Personality and Individual Differences*, *169*, 110066. https://doi.org/10.1016/j.paid.2020.110066

Buecker, S., Denissen, J. J. A., & Luhmann, M. (2021). A propensity-score matched study of changes in loneliness surrounding major life events. *Journal of Personality and Social Psychology*, 121(3), 669-690. 10.1037/pspp0000373

Buecker, S., Maes, M., Denissen, J. J. A., & Luhmann, M. (2020). Loneliness and the Big Five Personality Traits: A Meta–Analysis. *European Journal of Personality*, *34*(1), 8-28. https://doi.org/10.1002/per.2229

Cacioppo, J.T., & Cacioppo, S. (2012). The phenotype of loneliness. *European Journal of Developmental Psychology*, *9* (4), 446-452. 10.1080/17405629.2012.690510

Cihangir, S., Barreto, M., & Ellemers, N. (2010). The dark side of ambiguous discrimination: How state self-esteem moderates emotional and behavioral responses to ambiguous and unambiguous discrimination. *British Journal of Social Psychology*, *49*(1), 155-174. 10.1348/014466609X425869

Dahlberg, L., McKee, K.J., Frank, A., & Naseer, M. (2021). A systematic review of longitudinal risk factors for loneliness in older adults. *Aging & Mental Health*, *26*(2), 225-249. https://doi.org/10.1080/13607863.2021.1876638

Department for Digital, Culture, Media, and Sport (2022). Tackling Loneliness Evidence Review. https://www.gov.uk/government/publications/tackling-loneliness-evidence-review/tackling-loneliness-evidence-review-full-report#authors

Denissen, J.J.A. & Penke, L. (2008). Neuroticism predicts reactions to cues of social inclusion. *European Journal of Personality, 22*, 497-517. doi: 10.1002/per.682

Doyle, D. M., & Molix, L. (2014). Love on the margins: The effects of social stigma and relationship length on romantic relationship quality. *Social Psychological and Personality Science*, *5*, 102–110. 10.1177/1948550613486677

Doyle, D.M., & Molix, L. (2016). Disparities in Social Health by Sexual Orientation and the Etiologic Role of Self-Reported Discrimination. *Archives of Sexual Behavior, 45*, 1317–1327. doi:10.1007/s10508-015-0639-5

Du, H., Li, X., Peilian, C., Zhao, S., & Zhao, J. (2018). Loneliness and self-esteem in children and adolescents affected by parental HIV: A 3-Year longitudinal study. *Applied Psychology: Health & Wellbeing, 11*(1), 3-19. https://doi.org/10.1111/aphw.12139

Eccles, A.M., & Qualter, P. (2020). Alleviating loneliness in young people: A meta-analysis of interventions. *Child and Adolescent Mental Health, 6*, 17-33. 10.1111/camh.12389

Ejlskov, L., Wulff, J., Boggild, H., Kuh, D., Stafford, M. (2018). Assessing the relative importance of correlates of loneliness in later life: Gaining insight using recursive partitioning. *Aging & Mental Health, 22*(11), 1486-1493. doi: 10.1080/13607863.2017.1370690

Friedman, J. H., (2001). Greedy function approximation: A gradient boosting machine. *The Annals of Statistics*, *29*, 1189-1232. 10.1214/aos/1013203451

Funk, J. L. & Rogge, R. D. (2007). Testing the ruler with item response theory: Increasing precision of measurement for relationship satisfaction with the Couples Satisfaction Index. *Journal of Family Psychology, 21*, 572-583. 10.1037/0893-3200.21.4.572

Gosling, S. D., Rentfrow, P. J., & Swann, W. B., Jr. (2003). A Very Brief Measure of the Big Five Personality Domains. *Journal of Research in Personality, 37, 504-528.* https://doi.org/10.1016/S0092-6566(03)00046-1.

Griffin, S. C., Williams, A. B., Ravyts, S. G., Mladen, S. N. & Rybarczyk, B. D. (2020). Loneliness and sleep: A systematic review and meta-analysis. Health Psychology Open 7(1), 2055102920913235. 10.1177/2055102920913235

Hawkley, L. C., Browne, M. W., & Cacioppo, J. T. (2005). How Can I Connect With Thee?: Let Me Count the Ways. *Psychological Science*, *16*(10), 798–804. doi:10.1111/j.1467-9280.2005.01617.x

Heu, L. C., van Zomeren, M., & Hansen, N. (2021). Does Loneliness Thrive in Relational Freedom or Restriction? The Culture-Loneliness Framework. Review of General Psychology, 25(1), 60-72. https://doi.org/10.1177/1089268020959033

Hofstede, G. (1991). *Cultures and organizations: Software of the mind*. USA: McGraw-Hill.

Holt-Lunstad, J., Robles, T. F., & Sbarra, D. A. (2017). Advancing social connection as a public health priority in the United States. *The American Psychologist*, 72(6), 517-530. doi: 10.1037/amp0000103

James, G., Witten, D., Hastie, T., & Tibshirani, R. (2013). *An introduction to statistical learning.* Springer. https://doi.org/10.1007/978-1-4614-7138-7

Jefferson, R., Barreto, M., Jones, F., Conway., J., Chohan, A., Rich-Madsen, K., Verity, L., & Qualter, P. (2023a). Adolescent loneliness across the world and its relation to culture, school climate, and academic performance. *British Journal of Educational Psychology, 93(4), 997-1016.* 10.1111/bjep.12616

Jefferson, R., Barreto, M., Verity, L., Fit2Belong, Qualter, P. (2023b). Loneliness during the school years: How it affects learning and how schools can help. *Journal of School Health, 93*(5), 428-435. https://doi.org/10.1111/josh.13306

Joseph, V. R. (2022). Optimal ratio for data splitting. *Statistical Analysis and Data Mining: The ASA Data Science Journal*, *15*, 531–538. https://doi.org/10.1002/sam.11583

Kidd, S.A. (2007). Youth Homelessness and Social Stigma. *Journal of Youth and Adolescence*, *36*, 291–299. doi:10.1007/s10964-006-9100-3

Kleftaras, G., & Vasilou, E. (2016) Spirituality and the psychological impact of unemployment: Personality characteristics, loneliness, and depressive symptomatology. *The European Journal of Counselling Psychology*, *5*(1), 62-77. https://doi.org/10.5964/ejcop.v5i1.90

Kung, C. S. J., Kunz, J. S. & Shields, M. A. (2021). Economic aspects of loneliness in Australia. *Australian Economic Review, 54*, 147–163; doi: 10.1111/1467-8462.12414

Kyriazos, T., Galanakis, M., Karakasidou, E., & Stalikas, A (2021). Early COVID-19 quarantine: A machine learning approach to model what differentiated the top 25% well-being scorers. *Personality and Individual Differences, 181*. 10.1016/j.paid.2021.110980.

Lasgaard, M., Løvschall, C., Qualter, P., Mølgaard Laustsen, L., Lim, M.H., Terkildsen Maindal., H., Hargaard, A-S., Christiansen, J. (2022). Are loneliness interventions effective in reducing loneliness? A meta-analytic review of 128 studies. European Journal of Public Health, 32 (Supl 3). 10.1093/eurpub/ckac129.266

Lauder, W., Sharkey, S. and Mummery, K. (2004), A community survey of loneliness. *Journal of Advanced Nursing, 46*, 88-94. doi:10.1111/j.1365-2648.2003.02968.x

Lee, Y., & Bierman, A. (2019). Loneliness as a Mediator of Perceived Discrimination and Depression: Examining Education Contingencies. *The International Journal of Aging and Human Development, 89*(2), 206–227. doi:10.1177/0091415018763402

Luhmann, M., & Hawkley, L. C. (2016). Age differences in loneliness from late adolescence to oldest old age. *Developmental Psychology, 52*(6), 943-959. doi:10.1037/dev0000117

Madsen, K.R., et al., (2016). Loneliness, immigration background and self-identified ethnicity: a nationally representative study of adolescents in Denmark. *Journal of Ethnic and Migration Studies, 42*(12), 1977-1995, doi:10.1080/1369183X.2015.1137754

Maes, M., Qualter, P., Vanhalst, J., Van den Noortgate, & Goossens, L., (2019). Gender differences in loneliness across the lifespan: A meta-analysis. *European Journal of Personality, 33*, 642-654. https://doi.org/10.1002/per.2220

Mann, F., Bone, J.K., Lloyd-Evans, B., Frerichs, J., Pinfold, V., Ma, R., et al. (2017). A life less lonely: The state of the art in interventions to reduce loneliness in people with mental health problems. *Social Psychiatry and Psychiatric Epidemiology*, *52*(6), 627-38. doi:10.1007/s00127-017-1392-y

Martin, K. S., Rogers, B. L., Cook, J. T., & Joseph, H. M. (2004). Social capital is associated with decreased risk of hunger. *Social Science and Medicine*, *58*, 2645-2654. 10.1016/j.socscimed.2003.09.026

Matthews, T., Odgers, C. L., Danese, A., Fisher, H. L., Newbury, J. B., Caspi, A., Moffitt, T. E., & Arseneault, L. (2019). Loneliness and Neighborhood Characteristics: A Multi-Informant, Nationally Representative Study of Young Adults. *Psychological Science*, *30*, 765 –775. 10.1177/0956797619836102

Mihalopoulos, C. et al. (2020). The economic costs of loneliness: A review of cost-of-illness and economic evaluation studies. *Social Psychiatry and Psychiatric Epidemiology 55*, 823–836. doi: 10.1007/s00127-019-01733-7.

Morgan, K., Melendez-Torres, G.J., Bond, A., Hawkins, J, Hewitt, G., Murphy, S. & Moore, G. (2019). Socio-Economic Inequalities in Adolescent Summer Holiday Experiences, and Mental Wellbeing on Return to School: Analysis of the School Health Research Network/Health Behaviour in School-Aged Children Survey in Wales. *International Journal of Environmental Research and Public Health, 16*, 1107. 10.3390/ijerph16071107

Mund, M., & Johnson, M.D. (2021). Lonely Me, Lonely You: Loneliness and the Longitudinal Course of Relationship Satisfaction. *Journal of Happiness Studies*, *22*, 575–597. doi: 10.1007/s10902-020-00241-9

Murray, S. L., Rose, P., Bellavia, G. M., Holmes, J. G., & Kusche, A. G. (2002). When rejection stings: How self-esteem constrains relationship enhancement processes. *Journal of*

Personality and Social Psychology, *83*(3), 556–573. https://doi.org/10.1037/0022-3514.83.3.556

ONS. (2018). Loneliness - What characteristics and circumstances are associated with feeling lonely?https://www.ons.gov.uk/peoplepopulationandcommunity/wellbeing/articles/lonelinesswhatcharacteristicsandcircumstancesareassociatedwithfeelinglonely/2018-04-10

Park, C. et al. (2020). The effect of loneliness on distinct health outcomes: A comprehensive review and meta analysis. Psychiatry Research, 294. 10.1016/j.psychres.2020.113514.

Perlman, D. & Peplau, L. A. (1981). Towards a social psychology of loneliness. In S. Duck and R. Gilmour (Eds.) *Personal relationships in disorder*. London: Academic Press.

Priest, N., Perry, R., Ferdinand, A., Kelaher, M., & Paradies, Y. (2017). Effects over time of self-reported direct and vicarious racial discrimination on depressive symptoms and loneliness among Australian school students. *BMC Psychiatry*, *17*, 50. https://doi.org/10.1186/s12888-017-1216-3

Qin, Y., Ye, J., et al. (2022). Assessing the Nonlinear Effect of Atmospheric Variables on Primary and Oxygenated Organic Aerosol Concentration Using Machine Learning. *ACS Earth and Space Chemistry*, *6*(4), 1059-1066, doi:10.1021/acsearthspacechem.1c00443

Qualter, P., Vanhalst, J., Harris, R.A., Lodder, G., Bangee, M., Van Roekel, E., Verhagen, M., & Maes, M. (2015). Loneliness Across the Lifespan. *Perspectives on Psychological Science, 10*, 250-264, doi:10.1177/1745691615568999

Quan NG, Lohman MC, Resciniti NV, Friedman DB. (2019). A systematic review of interventions for loneliness among older adults living in long-term care facilities. *Aging Ment Health*, 1-11. doi:10.1080/13607863.2019.1673311

Rodriguez, M., Bellet, B.W. & McNally, R.J. (2020). Reframing Time Spent Alone: Reappraisal Buffers the Emotional Effects of Isolation. *Cognitive Therapy Research, 44*, 1052–1067. doi:10.1007/s10608-020-10128-x

Rosenberg, M. (1965). *Society and the adolescent self-image*. Princeton, NJ: Princeton University Press.

Russell, D. (1996). UCLA loneliness scale (version 3): Reliability, validity, and factor structure. *Journal of Personality Assessment*, *66*, 20-40. doi:/10.1207/s15327752jpa6601_2

Sternthal, M., Slopen, N. , Williams, D. R. (2011). Racial disparities in health: How much does stress really matter? *DuBois Review*, *8*(1), 95-113.

Thomson, R., Yuki, M., Talhelm, T., Schug, J., Kito, M., et al. (2018). Relational mobility predicts social behaviours in 30 countries and is tied to historical farming and threat. *Proceedings of the National Academy of Sciences, 115*(29), 7521-7526. https://doi.org/10.1073/pnas.1713191115

Tough, H., Siegrist, J., & Fekete, C. (2017). Social relationships, mental health, and wellbeing in physical disability: A systematic review. *BMC Public Health*, *17*, 414. https://doi.org/10.1186/s12889-017-4308-6

Triandis, H. (1995). *Individualism and collectivism*. NY: Routledge.

Victor, C., Scambler, S., Bond, J., & Bowling, A. (2000). Being alone in later life: Loneliness, social isolation and living alone. *Reviews in Clinical Gerontology, 10,* 407– 417. doi: 10.1017/S0959259800104101

Victor, C.R., Burholt, V. & Martin, W. (2012). Loneliness and Ethnic Minority Elders in Great Britain: An Exploratory Study. *Journal of Cross-Cultural Gerontology, 27,* 65–78. 10.1007/s10823-012-9161-6

Wang, J. M., Rubin, K. H., Laursen, B., Booth-LaForce, C., & Rose-Krasnor, L. (2013). Preference-for-solitude and adjustment difficulties in early and late adolescence. *Journal of Clinical Child and Adolescent Psychology*, *42*, 834–842. 10.1080/15374416.2013.794700.

Weiss, R. S. (1973). *Loneliness: The experience of emotional and social isolation*. The MIT Press.

WHO (2023). https://www.who.int/news/item/15-11-2023-who-launches-commission-to-foster-social-connection

Yuki, M., & Schug, J. (2020). Psychological consequences of relational mobility. *Current Opinion in Psychology*, *32*, 129-132. https://doi.org/10.1016/j.copsyc.2019.07.029.

Zautra, A.J., Affleck, G.G., Tennen, H., Reich, J.W. and Davis, M.C. (2005), Dynamic Approaches to Emotions and Stress in Everyday Life: Bolger and Zuckerman Reloaded With Positive as Well as Negative Affects. *Journal of Personality*, *73*, 1511-1538. doi:10.1111/j.0022-3506.2005.00357.x

Table 1

Table 1. Descriptive statistics for all variables included in the analyses.

Variable	Description	N (%)	Mean (std)
Loneliness frequency (UCLA mean)	Scale 1-5	-	2.66 (1.13)
Gender	Male	12811 (32%)	
	Female	27269 (68%)	
Age	16-24	2899 (7.2%)	
	25-34	5230 (13.0%)	
	35-44	6170 (15.4%)	
	45-54	9139 (22.8%)	
	55-64	9786 (24.4%)	
	65-74	5782 (14.4%)	
	75 +	1074 (2.7%)	
Employment status	Employed	37757 (94.2%)	
	Unemployed	2253 (5.6%)	
Years of education	<10 years	1422 (3.5%)	
	11 - 14 years	7197 (17.9%)	
	> 15 years	31461 (78.5%)	
Income	Poorly	6669 (16.6%)	
	Fairly well	19910 (49.7%)	
	Very well	13501 (33.7%)	
SES	Scale (1-10)		6.12 (1.81)
Live alone and choice	Alone and choose alone	24338 (60.7%)	-
	Alone but choose not to	6804 (17.0%)	
	Not alone and choose not to	8938 (22.3%)	

Live alone years	Open number	-	4.56 (11.12)
Number of people in household	Open number	-	1.23 (1.40)
Marital status	Single	11644 (29.0%)	-
	In a relationship but not living together	2295 (5.7%)	
	Married or Cohabiting	16463 (41%)	
	Divorced or Separated	7409 (18.5)	
	Widowed	2269 (5.7%)	
Sexual orientation	Exclusively Heterosexual	30849 (76.9%)	
	Predominantly Heterosexual	5051 (12.6%)	
	Equal	933 (2.3%)	
	Predominantly Homosexual	730 (1.8%)	
	Exclusively homosexual	1434 (3.5%)	
	Asexual	1083 (2.7%)	
Dependants	No dependant	28465 (71.0%)	-
	Have dependant	11615 (29.0%)	
Career length (years)	Open number		0.09 (0.40)
Age of youngest child (months)	Open number		136.41(176.15)
Number of children	Open number		1.04 (1.33)
Couple satisfaction	Scale (4 – 32)		16.56 (5.43)
Loneliness positive	Scale (1-3) No = 1; Sometimes = 2; Yes = 3		1.47 (0.56)
Hours spent alone	Open number		11.63 (7.20)
Enjoy time alone	Scale (1-5) Not at all = 1; Very much = 5		3.39(0.97)
Personality extraversion	Scale (1-7)		3.71 (1.49)
Personality agreeableness	Scale (1-7)		4.79 (1.25)
Personality conscientiousness	Scale (1-7)		5.29 (1.21)
Personality emotional stability	Scale (1-7)		4.51 (1.45)
Personality openness	Scale (1-7)		5.06 (1.23)
Experiences with prejudice	Scale (1-7)		2.36 (0.97)
Personal self esteem	Scale (4 – 32)		17.25 (3.13)
Social capital	Scale (1-5)		3.00 (0.73)
Relational mobility	Scale (1 – 7)		3.97 (0.85)
Residence in same country as birth	Same	27809 (69.4%)	
	Not the same	12271 (30.6%)	

Hofestede individualism	1-100	83.80 (14.92)
Hofestede power distance	1-100	38.43 (10.75)
Health self-rating	Scale (1-5)	3.41 (1.02)

Figure 1-3

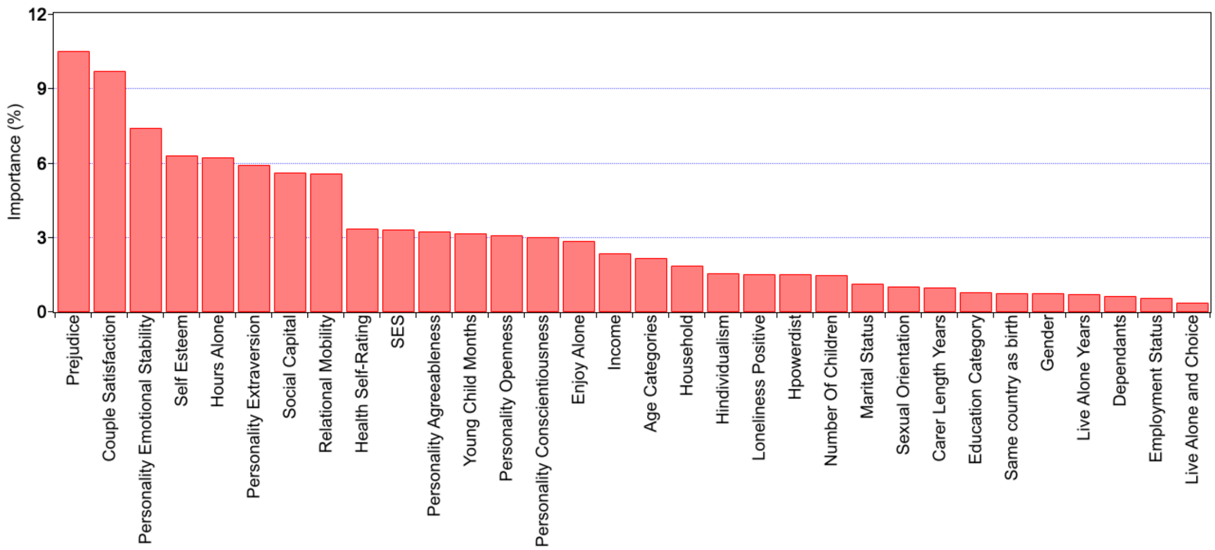

Figure 1. Importance of identifiers of loneliness frequency in random forests prediction

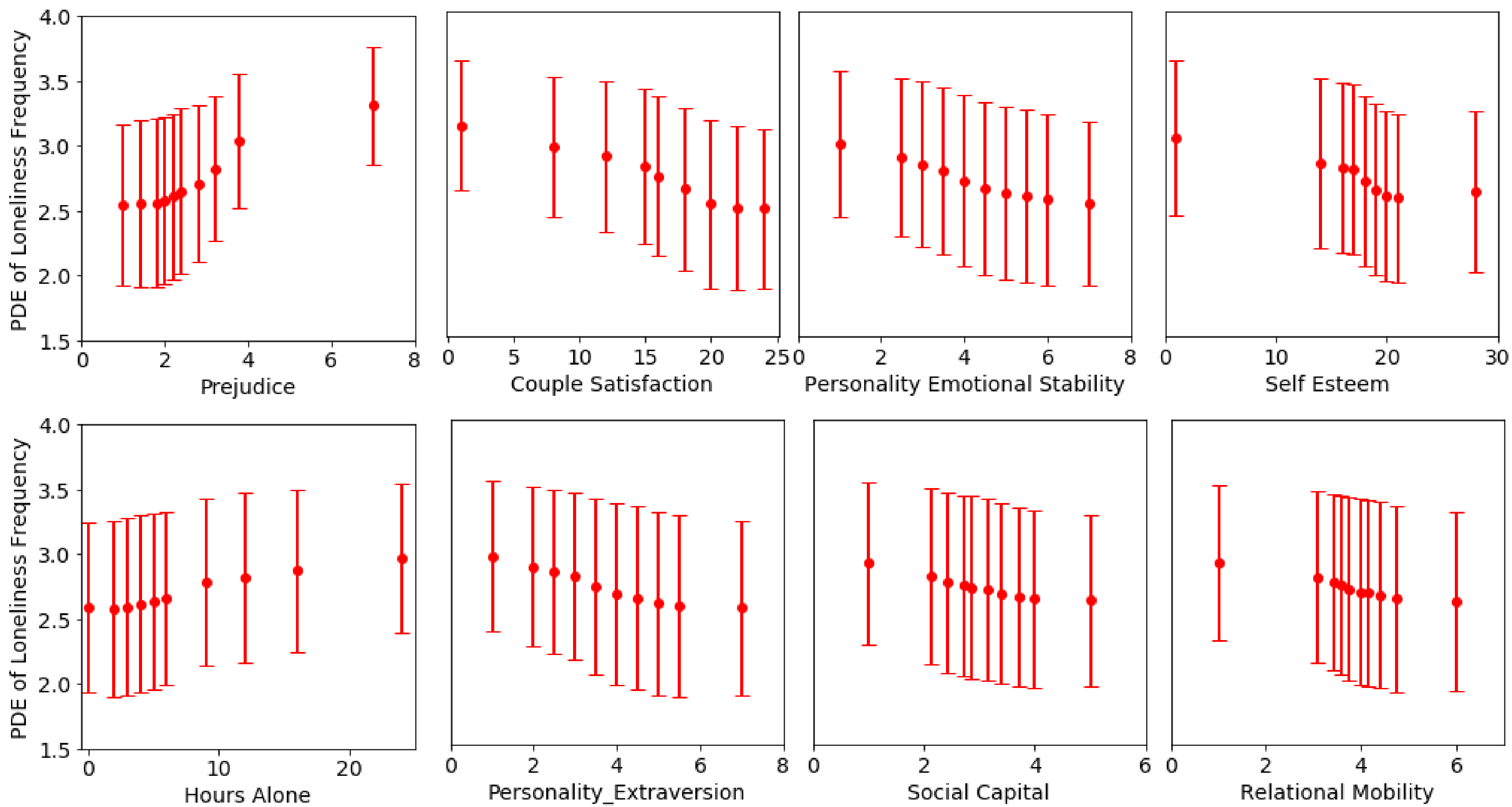

Figure 2. 1-D Partial dependence plots of loneliness frequency with respect to the eight most important features. The y-axis is the partial dependence of expected loneliness frequency as a function of the input features of interest. The solid dots represent the average estimation, and the error bars represent the standard deviation.

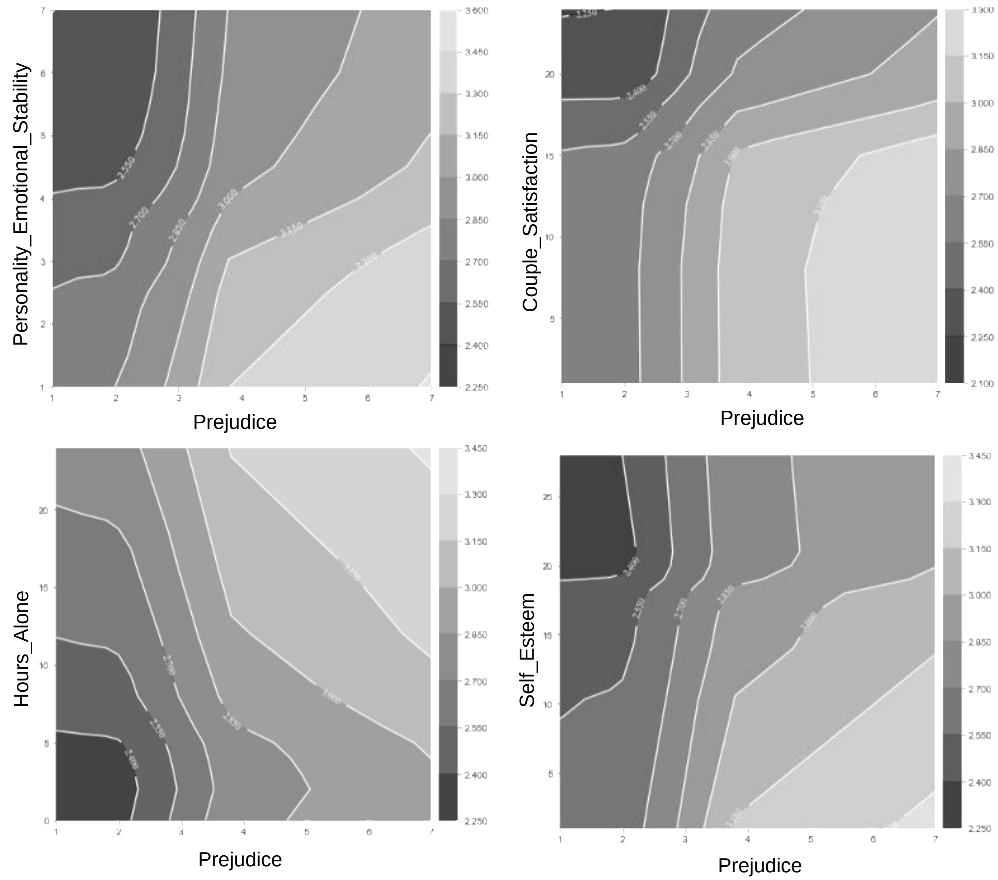

Figure 3. 2-D Partial dependence plots of loneliness frequency and the interactions of Prejudice with Emotional Stability, Couple Satisfaction, Hours Alone, and Personal Self Esteem. Colored contour bands represent ranges of loneliness frequency prediction. The light color represents high loneliness frequency, and the dark color represents low loneliness frequency. The exact loneliness frequency values in each contour bands are also listed in the figure.

Supplementary materials for:

Understanding the Psychological, Relational, Socio-Cultural, and Demographic Predictors of Loneliness Using Explainable Machine Learning

The supporting materials include:

1. **Text S1:** The interaction of prejudice, couple satisfaction, emotional stability, and extroversion, with age and gender for the loneliness frequency prediction.
2. **Figure S1.** 2-D Partial dependence plots of loneliness frequency and the interactions of Prejudice, couple satisfaction, emotional stability, and extroversion, with Age.
3. **Figure S2.** 2-D Partial dependence plots of loneliness frequency and the interactions of Prejudice, couple satisfaction, emotional stability, and extroversion, with Gender.

Text S1: The interaction of prejudice, couple satisfaction, emotional stability, and extroversion, with age and gender for the loneliness frequency prediction

Since age was not one of the most important predictors, it was not included in these PDPs, but given its importance in the loneliness literature we opted to additionally explore whether it interacted with prejudice (the most important predictor), couple satisfaction (a variable that can plausibly vary in importance across the lifespan), and the two personality variables that emerged as important predictors, i.e., emotional stability and extroversion (given that past work has suggested that the relationship between the big five and loneliness might be moderated by age, Buecker et al., 2020) (Figure S1). Similarly, although gender did not emerge as one of the most important predictors, since it has been suggested that while gender might have little effect on loneliness, it might change its drivers (Maes et al., 2019), we re-did these plots with gender instead of age as the interacting factor (Figure S2). Again, the plots suggest that the association of loneliness with prejudice, couple satisfaction, emotional stability and extroversion is similar for men and women.

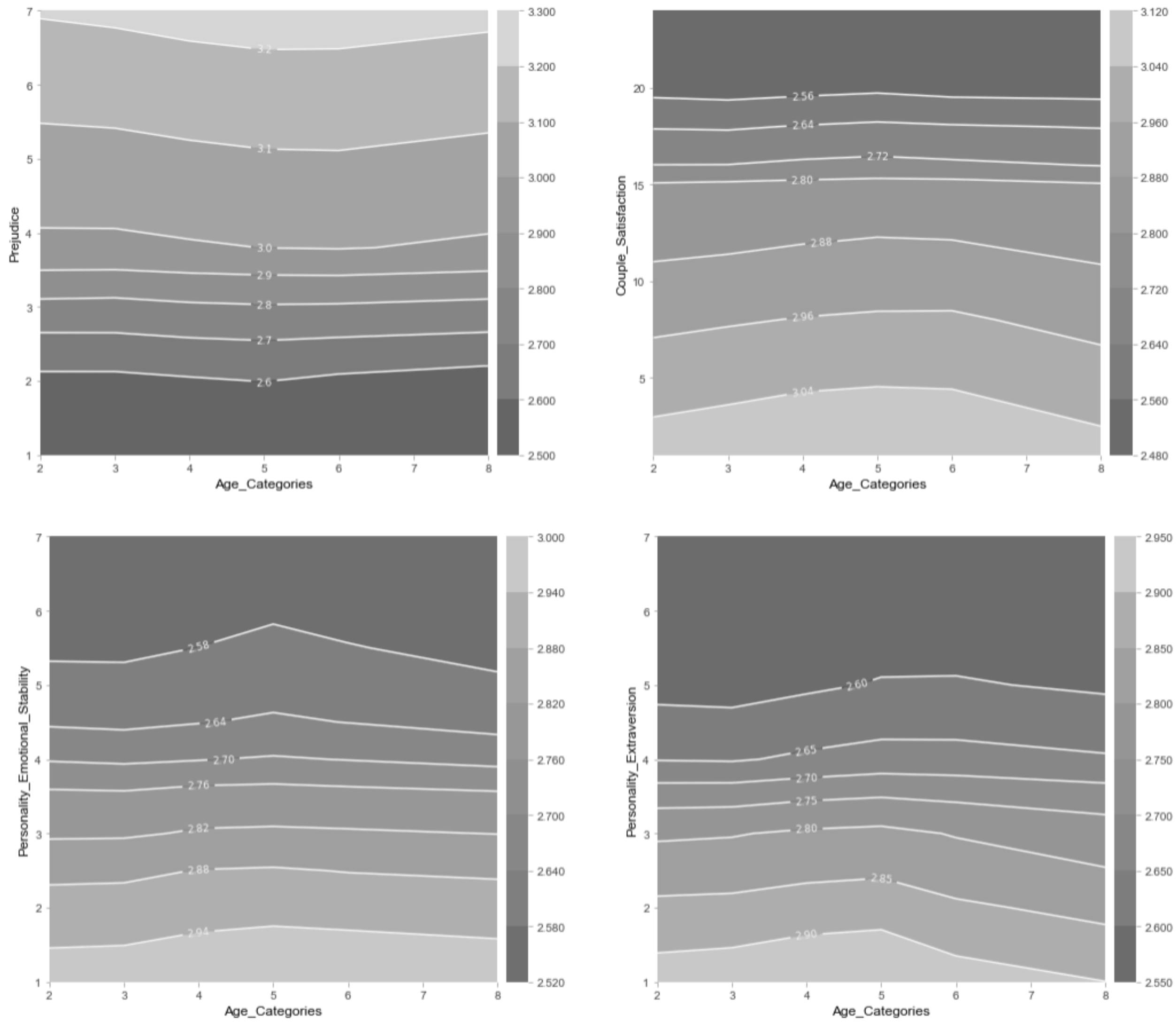

Figure S1. 2-D Partial dependence plots of loneliness frequency and the interactions of Prejudice, couple satisfaction, emotional stability, and extroversion, with Age. Colored contour bands represent ranges of loneliness frequency prediction. The light color represents high loneliness frequency, and the dark color represents low loneliness frequency. The exact loneliness frequency values in each contour bands are also listed as the color bar.

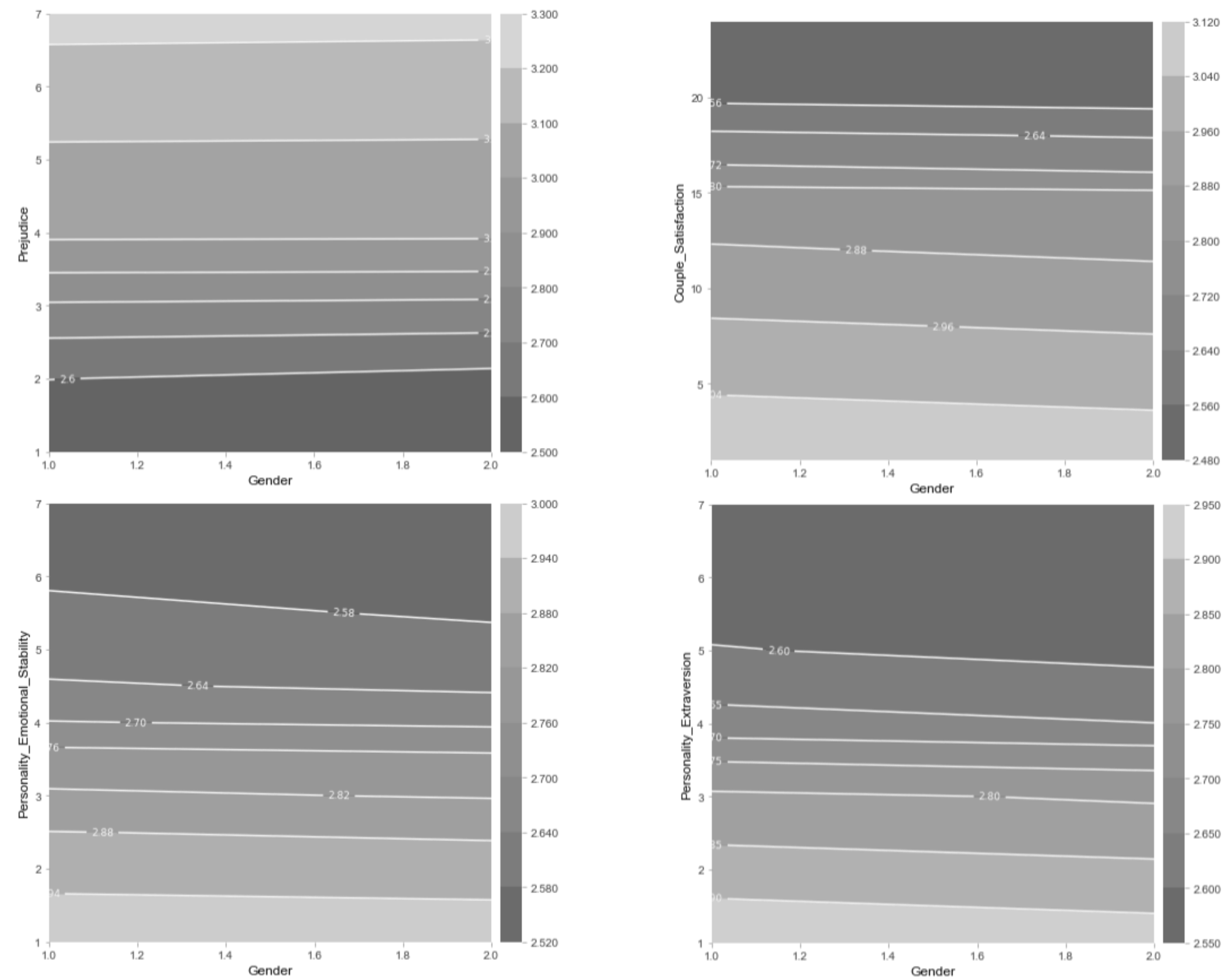

Figure S2. 2-D Partial dependence plots of loneliness frequency and the interactions of Prejudice, couple satisfaction, emotional stability, and extroversion, with Gender. Colored contour bands represent ranges of loneliness frequency prediction. The light color represents high loneliness frequency, and the dark color represents low loneliness frequency. The exact loneliness frequency values in each contour bands are also listed as the color bar.

References

Buecker, S., Maes, M., Denissen, J. J. A., & Luhmann, M. (2020). Loneliness and the Big Five Personality Traits: A Meta–Analysis. *European Journal of Personality, 34*(1), 8-28. https://doi.org/10.1002/per.2229

Maes, M., Qualter, P., Vanhalst, J., Van den Noortgate, & Goossens, L., (2019). Gender differences in loneliness across the lifespan: A meta-analysis. *European Journal of Personality, 33*, 642-654. https://doi.org/10.1002/per.2220