

Human Values and Digital Patterns in Physical Exercise (Extended Abstract)*

Yelena Mejova and Kyriaki Kalimeri

ISI Foundation, Turin, Italy
yelenamejova@acm.org,kalimeri@ieee.org

Abstract

In this study, we present a unique demographically representative dataset of 15k US residents that combines technology use logs with surveys on moral views, human values, and emotional contagion. First, we show which values determine the adoption of Health & Fitness mobile applications, finding that users who prioritize the value of *purity* and de-emphasize values of *conformity*, *hedonism*, and *security* are more likely to use such apps. Further, we achieve a weighted AUROC of .673 in predicting whether individual exercises and find a strong link of exercise to respondent socioeconomic status, as well as the value of *loyalty*.

1 Introduction

Over the last half-century, the daily occupation-related energy expenditure of US workers has decreased by more than 100 calories, with the proportion of jobs requiring at least moderate intensity physical activity declining from 48% to 20% in 2008 [Church *et al.*, 2011]. Recently, both entrepreneurs and public health officials have been attempting to use new technology to encourage physical activity. Latest research shows a complex relationship between psychology, technology use, and exercise, finding that increased physical activity associated with health app use is related to feelings of self-efficacy [Litman *et al.*, 2015], with yet other insights linking exercise to being extroverted, neurotic, and less agreeable [Hausenblas and Giacobbi Jr, 2004; Bruner *et al.*, 2013], as well as having implications for mental health [Rhodes and Smith, 2006].

This study is a unique view of the interaction between technology use, demographics, and value systems of a representative US population sample (originally in [Mejova and Kalimeri, 2019]). Just over 15k participants filled in questionnaires including the following psychometric measures: Moral Foundations [Haidt and Joseph, 2004], Schwartz Basic Human Values [Schwartz, 2012], and Emotional Contagion [Doherty, 1997]. Along with these, 5,008 respondents agreed to

allow the capture of their desktop browsing data, whereas another 2,625 allowed to capture their mobile app usage. Using this data, we contribute the following insights on psychological markers of health app use and the actual exercise behavior by: (1) Modeling “Health & Fitness” application use in relation to psychometric and demographic variables, (2) Predicting the engagement in physical exercise via the above variables, as well as browsing and application use data, (3) Revealing determinants of exercise among the types of variable, and (4) Comparing the exercise behaviors across mobile applications.

2 Related Work

Recently, to understand the impact of technological interventions for increasing physical activity, researchers examined the role of individual characteristics, attitudes, and lifestyle of users [Carroll *et al.*, 2017], demographic attributes such as gender, age, socioeconomic factors, and technology literacy [Ernsting *et al.*, 2017; Carroll *et al.*, 2017]. For instance, using self-reported technology use, [Ernsting *et al.*, 2017; Carroll *et al.*, 2017] found that two-thirds of their participants were using a smartphone, and those who engaged in Health & Fitness apps were more motivated to change or maintain a healthy lifestyle. A further association was confirmed between smartphone use and health literacy [Bailey *et al.*, 2015], and an association with age, with seniors (65 years and older).

However, human values are also known to influence people’s actions, but has so far been largely unexplored [Kerner and Goodyear, 2017; McLennan and Thompson, 2015]. For example, Lee *et al.* [Lee *et al.*, 2008] examined the value-expressive function of attitudes and achievement goal theory in predicting the moral attitudes of young athletes. Ball *et al.* [Ball *et al.*, 2010] studied the individual preferences for social support according to the values system in following a healthy lifestyle. Among works closest to our intention, Lathia *et al.* [Lathia *et al.*, 2017] assessed the relationship between physical activity and happiness via a smartphone app concluding that people that exercise more are happier.

3 Data Collection

Data presented in this study spans 15,021 subjects in the United States of America, selected using probabilistic, rep-

*Full version of the paper was originally published at the ACM Conference on User Modelling, Adaptation and Personalization (UMAP), 2019 [Mejova & Kalimeri, 2019].

representative sampling methodology, all of whom were incentivised to participate. After receiving informed consent from all participants for the collection, storage, and analysis of the data, as well as the acceptance of the privacy policy¹, we administered a series of questionnaires to gather demographic and psychometric data. Also, we asked the participants for the access to either their basic mobile or desktop data for one month, resulting in desktop activity data for 5,008 people (2,823 women) and mobile activity data for 2,625 people (1,544 women). The latter subset with activity data has been discussed in [Kalimeri *et al.*, 2018]².

Demographics. Basic demographic factors (age, gender, ethnicity), geographic factors (home location, expressed at the zip code level), socioeconomic factors (educational level, marital status, parenthood, wealth, income), health-related factors (exercise, smoke, and weight issues) and political orientation.

Psychometric: Moral Foundations. We employ the Moral Foundation Theory [Haidt and Graham, 2007] which we operationalized via the Moral Foundations Questionnaire (MFQ) [Graham *et al.*, 2011], a validated measure of the degree to which individuals endorse each of five dimensions: care/harm, fairness/cheating, loyalty/betrayal, authority/subversion, and purity/degradation.

Psychometric: Schwartz Basic Human Values. We employ the Portrait Values Questionnaire [Schwartz, 2012]. The questionnaire is based on self-assessments resulting in a numerical value per person for each of the ten basic values: self-direction, stimulation, hedonism, achievement, power, security, conformity, tradition, benevolence, and universalism. The above ten values can be clustered into four higher order values, so-called quadrant values and into two dimensions, as the sum of the individual items of which they consist: *Openness to change* (self-direction, stimulation) vs. *Conservation* (security, conformity, tradition) and *Self-enhancement* (universalism, benevolence) vs. *Self-transcendence* (power, achievement).

Psychometric: Emotional Contagion. We employ the well-established emotional contagion scale (EC) [Doherty, 1997] which assesses mimetic tendency to five basic emotions (love, happiness, fear, anger, and sadness), measuring the individual differences in susceptibility to “catching” and empathizing the emotions of others.

Digital Data: Desktop Browsing Data. For the participants who permitted the logging of their desktops’ web browsing data, 5,008 in total, we capture: (i) the domain names, and (ii) the average time spent online and (iii) the number of visits per day on each domain. All this information is aggregated by day, and only the domain names (and not the page or section of the websites) are stored, to ensure the privacy of the participants. Users with fewer than $N = 30$ unique domains are discarded.

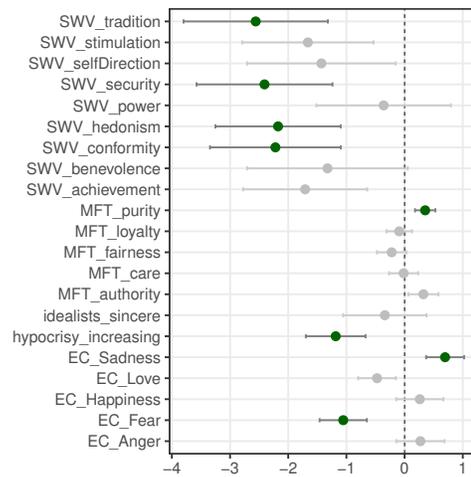


Figure 1: Coefficients of linear model predicting Health & Fitness app use, with those significant at $p < 0.05$ bolded in green.

Digital Data: Mobile Data. Participants are also asked to download an application which, upon agreement with the privacy policy, logs their web browsing activity and application usage, and 2,625 agreed to be tracked. Application usage was captured whenever the application was running in the “foreground”, including records of the date and time stamp, the local time zone, and time spent on the application (in seconds) and application’s Google Play Store category label. For mobile browsing data, URL data was captured from the native browser on the subject’s device (not any 3rd party browsers), only the URL domain was stored in consideration of privacy. Similar to the desktop browsing data, users with a number of visits fewer than $N = 30$ unique domains are discarded from the analysis leaving us with a total of 2,406 participants.

4 Health App Use

We begin by examining the usage of mobile applications (apps) in the Health & Fitness category. The only demographic variable associated with getting such an app (more precisely, opening it at least once in the time of observation), is gender, with females 45% more likely to get one than males. Apps marketed for tracking running activity (such as *RunKeeper* and *MapMyRun*) are 24.9% more likely to be adopted by females, whereas those for walking (*Walkroid* and *MapMyWalk*) are 101% more likely (that is, twice as likely). The distinction is even greater for weight loss applications, with females 168% more likely to adopt one than males.

Considering the psychometric attributes, we run a linear model ($n = 2620$) to predict the adoption of any Health & Fitness application, with the coefficients plotted in Figure 1, whiskers marking 95% confidence intervals, and those significant at $p < 0.05$ bolded in green. We find a negative relationship with values associated with *tradition*, *security*, *hedonism*, and *conformity*, as well as with concerns about *hypocrisy* increasing in the society, and a positive relationship with *purity* value. Emotional contagion results show a positive relationship with *sadness* but a negative with *fear*.

The top 30 apps by the proportion of respondents reporting exercising regularly are shown in Figure 2. Running, cycling, and walking tracking applications dominate the top, as well

¹<https://www.researchnow.com/privacy-policy/>

²For privacy considerations, the data will be made available upon request, exclusively for the scientific community.

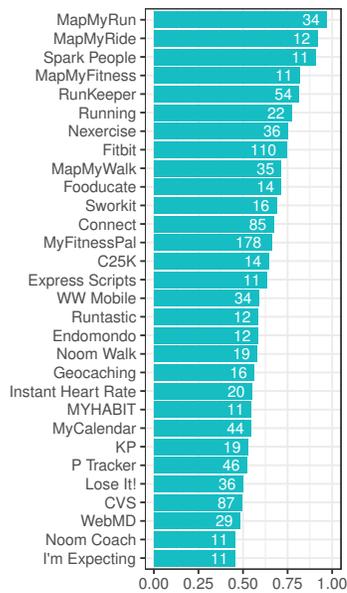


Figure 2: Applications in the Health & Fitness category ordered by the proportion of respondents reporting exercising regularly, with the n shown in white.

as *Spark People*, which provides a combination of weight loss and fitness (although the proportion is based on eleven respondents). Towards the bottom, we find generic health resources like *WebMD*, as well as pregnancy apps (*I'm Expecting*).

5 Predicting Exercise

Next, we would like to determine whether it is possible to use this data to predict engagement in physical activity. We formulate this study as a supervised classification problem, aiming at predicting whether participants exercise.

Focusing only on our “Mobile” dataset for which we have all the above information ($n=2620$), we train a Random Forest (RF) classifier [Breiman, 2001] inferring each time from a richer set of predictors starting from demographics, values/morals, health-related variables, web domain categories, application categories, and rate of usage of Health & Fitness applications. We perform ten-fold cross-validation and report the average Area Under the Receiver Operating Characteristic Curve (AUROC) weighted statistic over all folds.

We begin by attempting to predict exercise inferring only on the most basic demographic attributes – gender and age, resulting in AUROC of .513. However, adding richer demographic attributes, such as wealth, income, and educational level, significantly improved our prediction to 0.616 (at $p < 0.001$). Enhancing the baseline model with information about the moral values, we note an improvement in the performance (AUROC = .623). The increase is even more pronounced - as expected - with the inclusion of the health-related variables (.650). Adding web browsing domain categories (.654) slightly improved the model, but it is when including the categories of the applications used that the accuracy is increased in a notable way to .671. Including both sets of features - apps and domains categories - (.672), as

well as the average time people used a Health & Fitness application (.673) performs statistically identical to the previous model, indicating that the mere knowledge of application being opened once is enough.

5.1 Determinants of Exercise

In the aim of understanding the contribution of individual variables to whether a person exercises, we employ multivariate logistic regression analysis. The resulting models are shown in Table 1. In consideration of space, only variables which have a coefficient significant at $p < 0.01$ are shown; note however that the p-value markers (stars) shown in the table have been Bonferroni-adjusted to ameliorate the multiple comparison problems.

In the first model we find a strong positive relationship between exercise and *education*, *wealth*, and somewhat *income*, and a negative one with *age*. Considering the moral values, we find exercise to be related with *loyalty* and *stimulation*, and negatively related to *hedonism*, *power*, *security*, and *tradition*. Next, the addition of health-related variables unsurprisingly produces highly significant coefficients. There is a robust negative relationship between exercise and having *high blood pressure*, some other *chronic disease*, and being a *smoker*. Interestingly, we also find strong effects in the belief statements of respondents. Those who exercise are more likely to agree with the statement “When I think of making plans for the future, my health is something I strongly consider” (HQ 1), “We all have a choice about how to lead our lives, and healthy habits are just one example of that” (HQ 4), but are more likely to disagree with “Health is a gift and there is not much I can do about it” (HQ 5) and “Sometimes I avoid getting my test results if I think it will be bad news” (HQ 6). These findings underscore the importance of the individual’s belief in their ability to achieve goals [Litman *et al.*, 2015].

Note that for an intermediate model, not shown here for brevity, we add the internet browsing data available for both Desktop and Mobile cohorts. We find the most important domains to be in the areas of *Personal Network Storage* (content management), *Malicious Sites* (may include adult content), and *Interactive Web Applications* (document readers, calendar), though none had a particularly significant p-value after Bonferroni correction. Finally, as we consider the use of applications (those under `total_app.time`) for the respondents who shared their mobile activity, we find several application classes beneficial to the model, the most significant of these being *Health & Fitness*.

As the models increase in complexity (see the last two columns of Table 1), the Bonferroni correction becomes more strict, which reveals the variables most important in modeling exercise: a combination of demographics (wealth), attitude (HQ4 “Health is a choice”), and technology use (Health & Fitness Apps).

6 Discussion & Conclusions

This study is a contribution to an exciting area of research into the user-centered design of persuasive technologies and behavior change interventions to improve health and well-being. Here, we show that *values* held by the individuals may be an important dimension to consider. In a sense, it can be

seen as a response to a recent study on strategies to encourage diet and physical exercise by Radha et al. [Radha *et al.*, 2016], who called for the study of factors that could “explain attitude towards the feasibility level of a recommendation”.

In particular, insights obtained in this study lead us to recommend the following considerations when designing persuasive technologies for health behavior modification:

- Focus on a **particular activity** to track. We show that applications having most users exercising are marketed for tracking a particular activity, such as walking, running, or cycling. Those centered around particular wearables, for instance, fare less well.
- Create interventions with **socioeconomic status** of the users in mind. We find that wealth is one of the greatest determinants of regular exercise. More should be done to understand the barriers of the less wealthy to leading a healthy lifestyle. For instance, it is curious that it is wealth, not income, that remains most predictive in the regression model.
- Incorporate discussion of **values** in the interaction or interface. One of the strongest predictors for exercise is the belief that healthy habits are a conscious choice. Making this choice explicit may reinforce this value and encourage engagement.
- Encourage the expression of **happiness** and offer emotional rewards. We find people engaging in exercise identifying more strongly with the value of happiness, and although the direction of causation may point both ways, associating positive emotions with physical activity may reinforce the connection. Note that our findings that the value of *power* is negatively associated with exercise suggest that competitions and leader boards may not be appropriate for many users.
- Use application usage in **predictive analytics**, in the absence of detailed demographic or value information. In our classification experiments, we find mobile application usage to be much more useful in predicting exercise than internet browsing.

The greatest limitation of this study is the reliance on self-reporting when measuring exercise, as many biases are possible. In future studies, we encourage researchers also to obtain permission to gather user physical activity. Also, although we capture one month of technology use, the study is not longitudinal – many people may have downloaded and used the apps at some time, but not in our window of observation. Thus conclusions on application adoption/retention should be made while considering the small eventual sample size per application.

Finally, we would like to reiterate the privacy precautions taken in this study, with respondent anonymization, data aggregation, and URL cleaning, which is performed to limit the exposure of participants.

Acknowledgements

Y.M. and K.K., acknowledge support from the “Lagrange Project” of the Institute for Scientific Interchange (ISI) funded by the Fondazione Cassa di Risparmio di Torino (CRT).

	D+V+H (N+P+M)	D+V+H+ +U+A (M)
	n=15021 $R^2_{MF}=0.074$	n=2260 $R^2_{MF}=0.875$
(Intercept)	0.193	1.041
education	0.116 ***	0.124
gender	-0.082	0.098
income	0.044 *	0.070
parent	-0.109	-0.414
wealth	0.152 ***	0.174 **
age	-0.071 ***	-0.013
marital_status_married	-0.079	-0.123
political_party_vote	0.030	0.051
EC_Happiness	0.016	0.033
EC_Sadness	-0.021	-0.069
MFT_authority	-0.017	-0.011
MFT_loyalty	0.024 ***	-0.003
SWV_achievement	-0.059	-0.078
SWV_benevolence	-0.144 **	-0.148
SWV_hedonism	-0.130 ***	-0.167
SWV_power	-0.108 **	-0.146
SWV_security	-0.211 ***	-0.169
SWV_selfDirection	-0.071	0.023
SWV_stimulation	0.122 ***	0.170
SWV_tradition	-0.137 ***	-0.198
hypocrisy_increasing	-0.030	-0.058
blood_pressure_high	-0.424 ***	-0.549 *
chronic_disease	-0.465 ***	-0.326
smoker	-0.227 **	-0.142
Q17_1_health_plans	-0.158 ***	-0.085
Q17_4_habit_choice	-0.255 ***	-0.449 ***
Q17_5_health_is_gift	0.060 **	0.033
Q17_6_avoid_test_results	0.058 **	0.014
Education_Reference		0.003
General_News		-0.003
Interactive_Web_Apps		1.166
Internet_Radio_TV		0.112
Malicious_Sites		0.010
Motor_Vehicles		-0.035
Online_Shopping		-0.002
Personal_Net_Storage		0.291
Personal_Pages		-0.362
Recreation_Hobbies		0.023
Search_Engines		-0.002
Sports		0.012
total_web_visits		0.001
total_app_time		0.000
Entertainment		0.001
Health & Fitness		0.010 ***
House & Home		-0.011
Libraries & Demo		-0.540
Lifestyle		0.005
Maps & Navigation		0.005
Medical		-0.016
Music & Audio		0.004

Table 1: Logistic regression models predicting exercise using demographic (D), values (V), health (H), url domain (U), and app (A) features, applied to users who shared PC activity (P), mobile activity (M) or neither (N). Confidence levels: $p < 0.001$ ***, $p < 0.01$ **, $p < 0.05$ *.

References

- [Bailey *et al.*, 2015] Stacy C Bailey, Rachel O’conor, Elizabeth A Bojarski, Rebecca Mullen, Rachel E Patzer, Daniel Vicencio, Kara L Jacobson, Ruth M Parker, and Michael S Wolf. Literacy disparities in patient access and health-related use of internet and mobile technologies. *Health Expectations*, 18(6):3079–3087, 2015.
- [Ball *et al.*, 2010] Kylie Ball, Robert W Jeffery, Gavin Abbott, Sarah A McNaughton, and David Crawford. Is healthy behavior contagious: associations of social norms with physical activity and healthy eating. *International Journal of Behavioral Nutrition and Physical Activity*, 7(1):86, 2010.
- [Breiman, 2001] Leo Breiman. Random forests. *Machine learning*, 45(1):5–32, 2001.
- [Brunes *et al.*, 2013] Audun Brunes, Liv Berit Augestad, and Sigridur Lara Gudmundsdottir. Personality, physical activity, and symptoms of anxiety and depression: the hunt study. *Social psychiatry and psychiatric epidemiology*, 48(5):745–756, 2013.
- [Carroll *et al.*, 2017] Jennifer K Carroll, Anne Moorhead, Raymond Bond, William G LeBlanc, Robert J Petrella, and Kevin Fiscella. Who uses mobile phone health apps and does use matter? a secondary data analytics approach. *Journal of medical Internet research*, 19(4), 2017.
- [Church *et al.*, 2011] Timothy S Church, Diana M Thomas, Catrine Tudor-Locke, Peter T Katzmarzyk, Conrad P Earnest, Ruben Q Rodarte, Corby K Martin, Steven N Blair, and Claude Bouchard. Trends over 5 decades in us occupation-related physical activity and their associations with obesity. *PloS one*, 6(5):e19657, 2011.
- [Doherty, 1997] R William Doherty. The emotional contagion scale: A measure of individual differences. *Journal of nonverbal Behavior*, 21(2):131–154, 1997.
- [Ernsting *et al.*, 2017] Clemens Ernsting, Stephan U Dombrowski, Monika Oedekoven, Julie LO, et al. Using smartphones and health apps to change and manage health behaviors: a population-based survey. *Journal of medical Internet research*, 19(4), 2017.
- [Graham *et al.*, 2011] Jesse Graham, Brian a Nosek, Jonathan Haidt, Ravi Iyer, Spassena Koleva, and Peter H Ditto. Mapping the moral domain. *Journal of personality and social psychology*, 101(2):366–85, August 2011.
- [Haidt and Graham, 2007] Jonathan Haidt and Jesse Graham. When morality opposes justice: Conservatives have moral intuitions that liberals may not recognize. *Social Justice Research*, 20(1):98–116, 2007.
- [Haidt and Joseph, 2004] Jonathan Haidt and Craig Joseph. Intuitive ethics: How innately prepared intuitions generate culturally variable virtues. *Daedalus*, 133(4):55–66, 2004.
- [Hausenblas and Giacobbi Jr, 2004] Heather A Hausenblas and Peter R Giacobbi Jr. Relationship between exercise dependence symptoms and personality. *Personality and Individual differences*, 36(6):1265–1273, 2004.
- [Kalimeri *et al.*, 2018] Kyriaki Kalimeri, Mariano G Beiró, Matteo Delfino, Robert Raleigh, and Ciro Cattuto. Predicting demographics, moral foundations, and human values from digital behaviours. *Computers in Human Behavior*, 2018.
- [Kerner and Goodyear, 2017] Charlotte Kerner and Victoria A Goodyear. The motivational impact of wearable healthy lifestyle technologies: a self-determination perspective on fitbits with adolescents. *American Journal of Health Education*, 48(5):287–297, 2017.
- [Lathia *et al.*, 2017] Neal Lathia, Gillian M. Sandstrom, Cecilia Mascolo, and Peter J. Rentfrow. Happier people live more active lives: Using smartphones to link happiness and physical activity. *PLOS ONE*, 12(1):1–13, 01 2017.
- [Lee *et al.*, 2008] Martin J Lee, Jean Whitehead, Nikos Ntoumanis, and Antonis Hatzigeorgiadis. Relationships among values, achievement orientations, and attitudes in youth sport. *Journal of sport and exercise psychology*, 30(5):588–610, 2008.
- [Litman *et al.*, 2015] Leib Litman, Zohn Rosen, David Spierer, Sarah Weinberger-Litman, Akiva Goldschein, and Jonathan Robinson. Mobile exercise apps and increased leisure time exercise activity: a moderated mediation analysis of the role of self-efficacy and barriers. *Journal of medical Internet research*, 17(8), 2015.
- [McLennan and Thompson, 2015] Nancy McLennan and Jannine Thompson. *Quality physical education (QPE): Guidelines for policy makers*. UNESCO Publishing, 2015.
- [Mejova and Kalimeri, 2019] Yelena Mejova and Kyriaki Kalimeri. Effect of values and technology use on exercise: Implications for personalized behavior change interventions. In *Proceedings of the 27th ACM Conference on User Modeling, Adaptation and Personalization*, pages 36–45, 2019.
- [Radha *et al.*, 2016] Mustafa Radha, Martijn C. Willemsen, Mark Boerhof, and Wijnand A. IJsselstein. Lifestyle recommendations for hypertension through rasch-based feasibility modeling. In *Proceedings of the 2016 Conference on User Modeling Adaptation and Personalization, UMAP ’16*, pages 239–247, New York, NY, USA, 2016. ACM.
- [Rhodes and Smith, 2006] RE Rhodes and NEI Smith. Personality correlates of physical activity: a review and meta-analysis. *British journal of sports medicine*, 40(12):958–965, 2006.
- [Schwartz, 2012] Shalom H Schwartz. An overview of the schwartz theory of basic values. *Online Readings in Psychology and Culture*, 2(1):11, 2012.