EÖTVÖS LORÁND UNIVERSITY
FACULTY OF EDUCATION AND PSYCHOLOGY

Nándor Hajdú

Exploring Contextual Determinants of Human Choice

Theses

Doctoral School of Psychology
Head of Doctoral School: prof. dr. Róbert Urbán

Behavioral Psychology Program
Head of the Program: prof. dr. Anna Veres-Székely

Supervisors: prof. dr. Balázs Zoltán Aczél, dr. Barnabás Imre Szászi

Budapest, 2024

Publications related to the dissertation:

Hajdú, N., Aczél, B., & Szászi, B. I. (2023). Extending the Choice Architecture Toolbox: The Choice Context Exploration. *SAGE OPEN*, *13*. doi:10.1177/21582440231216831

Hajdú, N., Schmidt, K., Ács, G., Röer, J. P., Mirisola, A., Giammusso, I., … Szászi, B. I. (2022). Contextual factors predicting compliance behavior during the COVID-19 pandemic. A machine learning analysis on survey data from 16 countries. *PLOS ONE*, *17*. doi:10.1371/journal.pone.0276970

Hajdú, N., Szászi, B. I., Aczél, B., & Nagy, T. (2024). Felügyelt gépi tanulási módszerek alkalmazása a pszichológiai kutatásokban. *MAGYAR PSZICHOLÓGIAI SZEMLE*. Retrieved from https://m2.mtmt.hu/api/publication/34784542

Szászi, B. I., Hajdú, N., Szécsi, P., Tipton, E., & Aczél, B. (2022). A machine learning analysis of the relationship of demographics and social gathering attendance from 41 countries during pandemic. *SCIENTIFIC REPORTS*, *12*. doi:10.1038/s41598-021-04305-5

Szaszi, B., Komandi, K., Hajdu, N., Tipton, E. (2022). Applying behavioral interventions in a new context. In.: Mažar, N., & Soman, D. (Eds.). (2022). *Behavioral Science in the Wild*. University of Toronto Press.

Other publications:

Buchanan, E. M., Lewis, S. C., Paris, B., Forscher, P. S., Pavlacic, J. M., Beshears, J. E., … Primbs, M. A. (2023). The Psychological Science Accelerator’s COVID-19 rapid-response dataset. *SCIENTIFIC DATA*, *10*. doi:10.1038/s41597-022-01811-7

Hoogeveen, S., Sarafoglou, A., Aczél, B., Aditya, Y., Alayan, A. J., Allen, P. J., … Wagenmakers, E.-J. (2023). A many-analysts approach to the relation between religiosity and well-being. *RELIGION BRAIN AND BEHAVIOUR*, *13*, 237–283. doi:10.1080/2153599X.2022.2070255

Hajdú, N., Szászi, B. I., & Aczél, B. (2020). *Extending the Choice Architecture Toolbox: The Choice Context Mapping*. Retrieved from https://m2.mtmt.hu/api/publication/32050582

Varga, D., Perecz, B., Fülöp, K., Sipos, A., Janszky, J. V., Hajdú, N., & Pál, E. (2023). Urinary titin in myotonic dystrophy type 1. *MUSCLE & NERVE*, *68*, 215–218. doi:10.1002/mus.27917

Dorison, C. A., Lerner, J. S., Heller, B. H., Rothman, A. J., Kawachi, I. I., Wang, K., … Coles, N. A. (2022). In COVID-19 Health Messaging, Loss Framing Increases Anxiety with Little-to-No Concomitant Benefits: Experimental Evidence from 84 Countries. *AFFECTIVE SCIENCE*, *3*, 577–602. doi:10.1007/s42761-022-00128-3

Varga, S., Horváth, J., Neszmélyi, B., & Hajdú, N. (2022). The emergence of action-effect-related motor adaptation amidst outcome unpredictability. *Program of the 22nd Conference of the European Society for Cognitive Psychology (ESCoP 2022)*, 63. Retrieved from https://m2.mtmt.hu/api/publication/33926365

Coles, N. A., March, D. S., Marmolejo-Ramos, F., Larsen, J. T., Arinze, N. C., Ndukaihe, I. L. G., … Liuzza, M. T. (2022). A multi-lab test of the facial feedback hypothesis by the Many Smiles Collaboration. *NATURE HUMAN BEHAVIOUR*, *6*, 1731–1742. doi:10.1038/s41562-022-01458-9

Legate, N., Ngyuen, T.-V., Weinstein, N., Moller, A., Legault, L., Vally, Z., … Primbs, M. A. (2022). A global experiment on motivating social distancing during the COVID-19 pandemic. *PROCEEDINGS OF THE NATIONAL ACADEMY OF SCIENCES OF THE UNITED STATES OF AMERICA*, *119*. doi:10.1073/pnas.2111091119

Varga, S., Neszmélyi, B., Hajdú, N., & Horváth, J. (2022). The emergence of action-effect-related motor adaptation amidst outcome unpredictability. *JOURNAL OF EXPERIMENTAL PSYCHOLOGY-HUMAN PERCEPTION AND PERFORMANCE*, *48*, 711–723. doi:10.1037/xhp0001021

Varga, S., Hajdú, N., Neszmélyi, B., György, E., & Horváth, J. (2021). Dynamic action-effect related motor adaptation. *5th International Conference of the European Society for Cognitive and Affective Neuroscience, ESCAN 2021*, 154. Retrieved from https://m2.mtmt.hu/api/publication/32775336

Jones, B. C., DeBruine, L. M., Flake, J. K., Liuzza, M. T., Antfolk, J., Arinze, N. C., … Coles, N. A. (2021). To which world regions does the valence-dominance model of social perception apply? *NATURE HUMAN BEHAVIOUR*, *5*, 159–169. doi:10.1038/s41562-020-01007-2

Lencsés, A., & Hajdú, N. (2017). *A tesztelési hatás vizsgálata érzelmi ingerek tanulása esetén*. Retrieved from https://m2.mtmt.hu/api/publication/32628776

Introduction

In recent years, the scientific community has grappled with the replication crisis, particularly within the field of psychology (Wicherts et al., 2016; Nosek et al., 2015). This crisis has prompted a reevaluation of research practices and methodologies, with a focus on improving theory construction and research validity. While various solutions have been proposed, such as promoting preregistration and replication, the underlying issue of insufficient theoretical frameworks persists. This dissertation aims to address this gap by advocating for the importance of exploratory research and demonstrating how machine learning methods can significantly enhance exploratory studies in psychology.

One of the key challenges in psychology, unlike fields such as theoretical physics, is the lack of a comprehensive program for theory development (Mischel, 2008; Borsboom et al., 2021). As noted by Mischel (2008), theories in psychology are often treated as individual entities rather than collaborative endeavors. This "toothbrush problem," where psychologists are reluctant to adopt theories formulated by others, has led to fragmentation and ambiguity in theory construction. While there may not be a shortage of theories, there is a pressing need for a concerted effort towards theory development.

Exploratory research plays a vital role in addressing this challenge by generating working theories from observed data patterns (De Groot & Spiekerman, 2020). Unlike confirmatory research, which tests hypotheses derived from existing theories, exploratory research aims to uncover new insights and patterns in the data. However, traditional statistical methods may not be sufficient to capture the complexities of behavioral phenomena. This is where machine learning methods offer a promising solution.

Machine learning, as defined by Jordan and Mitchell (2015), refers to data processing processes in which algorithms optimize models to make predictions. These methods hold tremendous potential for addressing the complexities inherent in behavioral science. By leveraging machine learning algorithms, researchers can extract patterns from complex datasets and uncover latent structures within behavioral phenomena. This computational approach offers a systematic framework for exploring the multidimensional nature of human behavior.

The advantages of machine learning over traditional statistical procedures are manifold. Firstly, machine learning can test research questions about the accuracy of prediction, providing insights into the underlying mechanisms of behavior (Yarkoni & Westfall, 2017). Secondly, machine learning methods can yield more accurate and robust estimates, enhancing the reliability of research findings. Thirdly, the results generated by machine learning models are often more accessible to laypeople and market actors, facilitating the practical application of psychological insights. Lastly, machine learning can inform theory building in ways beyond hypothesis testing, contributing to the development of comprehensive theoretical frameworks (Hajdu et al., 2023).

The replication crisis has underscored the need for methodological approaches that can accommodate the complexity of behavioral phenomena and facilitate theory building grounded in empirical evidence. By embracing the challenges posed by complexity and harnessing the power of machine learning, researchers have the opportunity to chart a new course for behavioral science—one characterized by rigor, transparency, and a deeper appreciation for the intricacies of human behavior. This dissertation aims to explore the application of machine learning methods in psychology, particularly in the context of behavioral interventions and choice architecture, to advance our understanding of psychological phenomena.

Addressing the replication crisis necessitates a deeper understanding of the theoretical foundations of psychology (Green, 2021; Scheel et al., 2021). While efforts have been made to promote preregistration and replication studies, the reality is that many hypotheses fail to survive thorough testing (Scheel et al., 2021). This suggests a fundamental issue with the existing theories in psychology. Unlike other scientific fields where theories are collaboratively developed, psychology often relies on theories formulated by individual researchers (Mischel, 2008). This decentralized approach impedes the development of cohesive theoretical frameworks and hinders progress in understanding human behavior.

The lack of robust theories in psychology has several implications. Firstly, it increases the risk of continually reinventing the wheel, as researchers struggle to establish connections between different phenomena (Kruglanski, 2001; Vallacher & Nowak, 1997). Secondly, it hampers efforts to develop effective interventions for psychological disorders, as theories are necessary for identifying causal relationships and guiding therapeutic approaches (Borsboom, 2017; Cramer et al., 2016). Thirdly, it limits the direction of future research, as studies often lack theoretical grounding and focus on exploring data without clear hypotheses (Van Lissa, 2022).

The emergence of open science initiatives has shed light on the prevalence of exploratory research in psychology (Van Lissa, 2022). Many studies unintentionally adopt an exploratory approach due to the lack of specificity in existing theories (Scheel, 2022). Confirmatory research, which relies on testing hypotheses derived from theories, requires strong and detailed theoretical frameworks (Szollosi & Donkin, 2021). However, most psychological theories lack these characteristics and are unable to generate specific, testable hypotheses. As a result, exploratory research has become a necessary tool for generating ideas and working theories from patterns observed in data.

Behavioral intervention research presents unique challenges compared to other fields of psychology (Bryan et al., 2021). Many intervention effects are context-dependent, necessitating a paradigm shift towards acknowledging and addressing heterogeneity in treatment effects. This paradigm shift requires increased attentiveness to the sources of heterogeneity, emphasis on measuring research context and sample characteristics, and the development of statistical methods to detect unexpected sources of heterogeneity.

Machine learning methods offer a promising approach to address the challenges of theory construction and exploratory research in psychology. By leveraging the computational power of machine learning algorithms, researchers can uncover patterns and relationships in complex datasets that traditional statistical methods may overlook (Agrawal et al., 2019). Moreover, machine learning techniques can facilitate hypothesis generation by identifying novel associations and predicting outcomes based on large-scale data patterns. This data-driven approach to theory building can complement traditional deductive methods, offering new insights and avenues for exploration in psychology.

In the following sections, this dissertation will delve into the application of machine learning methods in psychology, focusing on their utility in exploratory research and theory development. Our goal is to demonstrate how machine learning techniques can enhance our understanding of psychological phenomena and contribute to the advancement of theory in psychology. Additionally, we will explore the practical implications of incorporating machine learning into psychological research, particularly in the context of behavioral interventions and choice architecture. Through these investigations, we aim to highlight the potential of machine learning in psychology and inspire further exploration into its applications in theory construction and empirical research.

Title	Topic	Goal	Result	Published in
Extending the choice architecture toolbox: The Choice Context Exploration	Predictors of choosing stairs over elevators when going upstairs	To find the potential influencing factors of stairs-elevator choice in a setting in a specific environment, where most of the possible predictors could be measured	The selected influencing factors predicted choice with >90% accuracy	Sage Open
Contextual factors predicting compliance behavior during the COVID-19 pandemic: A machine learning analysis on survey data from 16 countries	Predictors of compliance with COVID-19 regulations	To find the potential influencing factors of compliance in a measurement configuration where the environment (the participant's home) could not be measured fully	The selected influencing factors predicted choice with 62 - 87% accuracy	PLoS ONE
A machine learning analysis of the relationship of demographics and social gathering attendance from 41 countries during	Demographic predictors of social gathering attendance	To find the potential influencing demographic factors of social gathering attendance, with no other contextual data available	The selected influencing factors predicted choice with 52 - 84% accuracy	Scientific Reports

pandemic				
Applying behavioral interventions in a new context	The role of heterogeneity in behavioral intervention planning	To aid experts in intervention planning raising awareness about heterogeneity	-	Behavioral Science in the Wild

Table 1. Articles and/or book chapters included in the dissertation with their respective main topic, goals, and results.

In the first research paper, we use machine learning tools to explore the reasons why people choose stairs over elevators when going up in a building, or vice versa. We looked for the potential explanatory variables that lead to the most accurate predictions of what people would do. In this situation, the decision was relatively easy to track, and every major theorized potentially influencing factor could be, and was accounted for. In our second research paper presented in this dissertation, we were looking for contextual factors predicting compliance behavior during the COVID-19 pandemic. In this case, the definition and the context of choice is more fuzzy than in the previous article, and fewer predictors were available. In our third paper, we were interested in the prediction of social gathering attendance during the COVID-19 pandemic, based on demographic variables. In this case, we had the least amount of information that could be used for prediction. Finally, after the three research papers, we present a book chapter that summarizes our thoughts on choice context exploration and gives insights on how to plan better interventions, accounting for contextual heterogeneity.

Chapter 1

In the realm of behavioral interventions, the effectiveness of nudges—subtle changes to choice architecture aimed at influencing behavior—has been a subject of considerable interest (Duflo et al., 2011; Silva & John, 2017). However, the success of nudges varies widely across different contexts, leading to questions about the factors that determine their efficacy (John et al., 2013; Brandon et al., 2017). This chapter argues that understanding the contextual factors influencing decision-making is essential for designing effective choice architecture interventions. To address this need, we present a procedural framework for detecting contextual influences.

We propose a systematic procedure, termed Choice Context Exploration, for identifying and assessing the contextual factors influencing decision-making. This procedure aims to provide researchers with a structured approach for exploring the multifaceted nature of choice contexts and understanding their impact on behavior. By systematically examining contextual influences, researchers can enhance the effectiveness and generalizability of behavioral interventions, ultimately contributing to the development of comprehensive theoretical frameworks for behavior change.

Context plays a pivotal role in shaping human behavior, encompassing a wide range of factors that influence decision-making (Rogers et al., 2020). While the concept of context is multifaceted, encompassing physical, intrapersonal, and sociocultural elements, its exploration is often overlooked in behavioral intervention research (Szaszi et al., 2018). Understanding the contextual factors at play is essential for designing tailored interventions that account for the diverse motivations and constraints individuals face in different situations.

Exploring contextual influences involves identifying the myriad factors that may impact decision-making and assessing their relative importance. This process requires collecting data from diverse sources, including experts, laypeople, and relevant literature, to create a comprehensive list of potential influencing factors. By systematically curating this information, researchers can gain insights into the complex interplay between context and behavior.

Choice Context Exploration

The Choice Context Exploration procedure comprises four key steps aimed at systematically exploring contextual influences on decision-making. Firstly, we collect potential influencing factors through surveys, interviews, and literature reviews, creating a comprehensive list of contextual variables. Secondly, we quantify the influence of these factors on behavior, using empirical data to estimate the strength and direction of their relationship with decision outcomes. Thirdly, we assess individuals' beliefs about contextual influences, exploring how perceptions align with empirical findings and identifying patterns among respondents. Finally, we conduct a comparative analysis to evaluate the consistency between beliefs and behavior, shedding light on the accuracy of individuals' insights into their decision-making processes.

In step 1, our goal was to gather a list of potential factors that might influence whether people use the stairs or the elevator. To achieve this, we surveyed a sample of university students and asked experts to provide open-ended responses about the potential factors. The research plan was approved by the local institutional ethical review board.

We randomly selected 500 individuals from the subject pool at our local university in Hungary, which consisted of students who had signed up for a course where they could participate in various studies in exchange for course credits. We recruited 392 individuals from the subject pool at our local university in Hungary, who received course credits as compensation. We asked these participants to list the contextual factors that they believed influenced their own and others' decisions between stairs and elevators. Secondly, we identified experts by compiling a list of those who had published at least one peer-reviewed research article on the topic of stair usage interventions in the past decade. We asked these experts to list the potential factors that might influence the choice between stairs and elevators. Out of the 47 experts we contacted, seven responded.

We then processed each of the collected responses, registering new categories for each type of influencing factor that was mentioned. If a newly processed response did not fit into any of the existing categories, a new category was created. Finally, we also reviewed relevant literature for additional contextual influencing factors: we searched for papers about interventions that targeted staircase and elevator use. As a result, 16 potential influencing factors were identified: *Appeal of stairs/elevator, Comfort/Laziness, Destination Floor, Elevator availability, Environmental consciousness, Fear of confined spaces and/or technical problems, Fatigue, Importance of Health/Sports, Luggage, Number of people in the elevator, Peer behavior, Physical limitations, Speed, Speed of elevator, Stairs/elevator physical availability, Temperature. Appeal* is a factor only suggested by experts and aggregates physical aspects of a staircase that make approaching it a better experience. An example of the mention of *Appeal* would be “the design of the stair should be inviting, open, bright, and ventilated”; another example is “the physical appearance and condition of the stairs (often not well lit, maybe smelly)”.

In Step 2, behavioral and contextual data were collected to evaluate the influence of factors identified in Step 1 on stair vs. elevator choice. The methods and analysis were pre-registered. Participants (n=523) were Hungarian university students who completed an online questionnaire over 10 days, reporting their choice between stairs and elevators and contextual factors influencing their decision. Factors identified in Step 1 were assessed using Likert-type scales or multiple-choice items. Participants were incentivized to report their behavior consistently over the study period. To examine the extent to which contextual factors influenced the choices made between stairs and elevators, we defined a mixed effect logistic regression model with the choice between stairs and elevators as the dependent variable and the measured contextual factors as independent variables. Visited buildings and IDs were treated as random effects. *Speed* and *Destination floor* were allowed to have varying slopes between different IDs, as it was plausible that these factors would have different effects on different individuals. We applied Lasso regularization to improve the interpretability and prediction accuracy of the regression models by selecting only a subset of variables, rather than using all of them, in the final model. *Temperature* and *Number of people waiting for the elevator* added the least amount of information, so their regression coefficients were penalized the most by the regularization process and were reduced to 0.

Next, we wanted to estimate how well the model explained the variation in individuals' choices. To do this, we calculated the squared correlation coefficient between the predicted values and the measured values, $R^2 = 0.76$, to estimate the variance in choosing the stairs or elevator explained by the model.

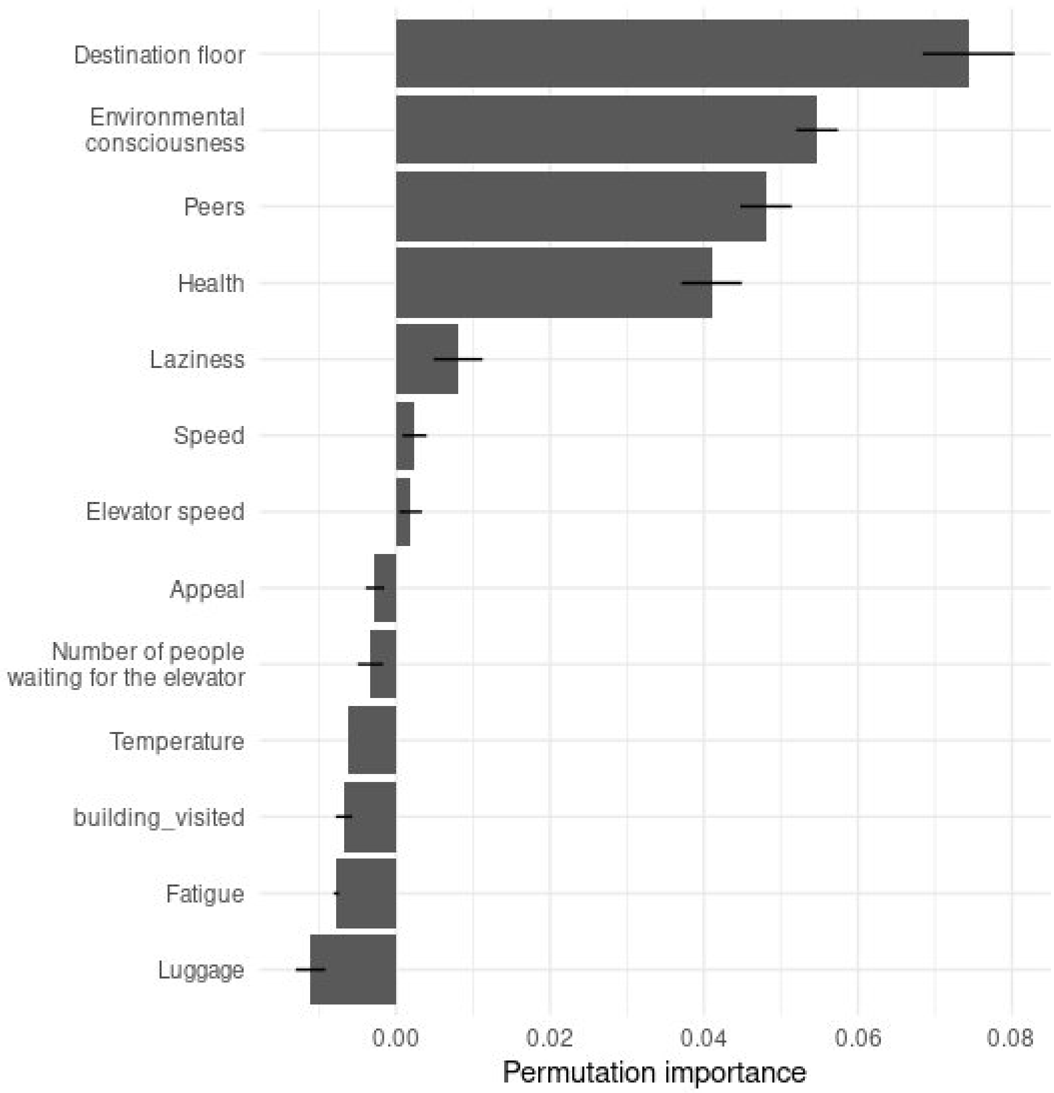

Figure 1. Permutation importance scores of predictors. The scores are the average decrease in the ROC AUC of 10 permutations, when the given variable was shuffled.

We also calculated permutation importance scores for each predictor to assess which of them were helpful for our model. The most important variables were *Destination floor, Environmental consciousness,* and *Peers*. Finally, we wanted to estimate the success of our model in correctly categorizing new data. We compared the model's predictions on the test data to the real decisions to assess the accuracy of the model. We used a probability threshold of .5, where predicted probabilities higher than .5 were categorized as someone choosing the stairs rather than the elevator. The results showed that the model correctly categorized 91.43% of the new cases.

Step 3 of the Choice Context Exploration aimed to gauge individuals' beliefs regarding the influence of contextual factors on their decisions between stairs and elevators, and to assess whether distinct groups with similar beliefs could be identified.

Methodology: Data was collected from 373 university students (298 female, 1 preferred not to answer) from the same subject pool used in Steps 1 and 2, with a mean age of 21.86 years (SD = 3.41), following approval from the local institutional ethical review board.

An online survey was devised to evaluate participants' beliefs concerning the importance of the contextual factors identified in Step 1. Participants rated the importance of each of the 13 factors identified in Step 2 using a Likert-type scale ranging from 0 (Not important at all) to 10 (Very important). The survey was completed either before or after the behavioral measurements conducted in Step 2. The survey instrument is accessible at https://osf.io/y7pmd/.

In order to explore individual differences regarding the factors influencing people's choices, we subjected the variables measuring the beliefs of participants to model-based clustering. The results show that participants can be divided into three groups in which its members hold similar beliefs about what influences their choices of stairs or elevators. Three clusters were defined and every cluster was named based on the pattern of factors. The first cluster, Efficiency group (N=118), had the highest mean scores on every scale except for Health and Environmental consciousness. Speed, Luggage, Laziness, Fatigue and Destination floor scales have the highest mean scores. The second cluster, the Health & Environment group (N = 68), had their highest mean scores on the Health and Environmental consciousness scales; every other mean score was low. In the third cluster, No priority group (N = 187), there was no substantial difference in the group mean scores, and between the group mean scores and the sample mean scores.

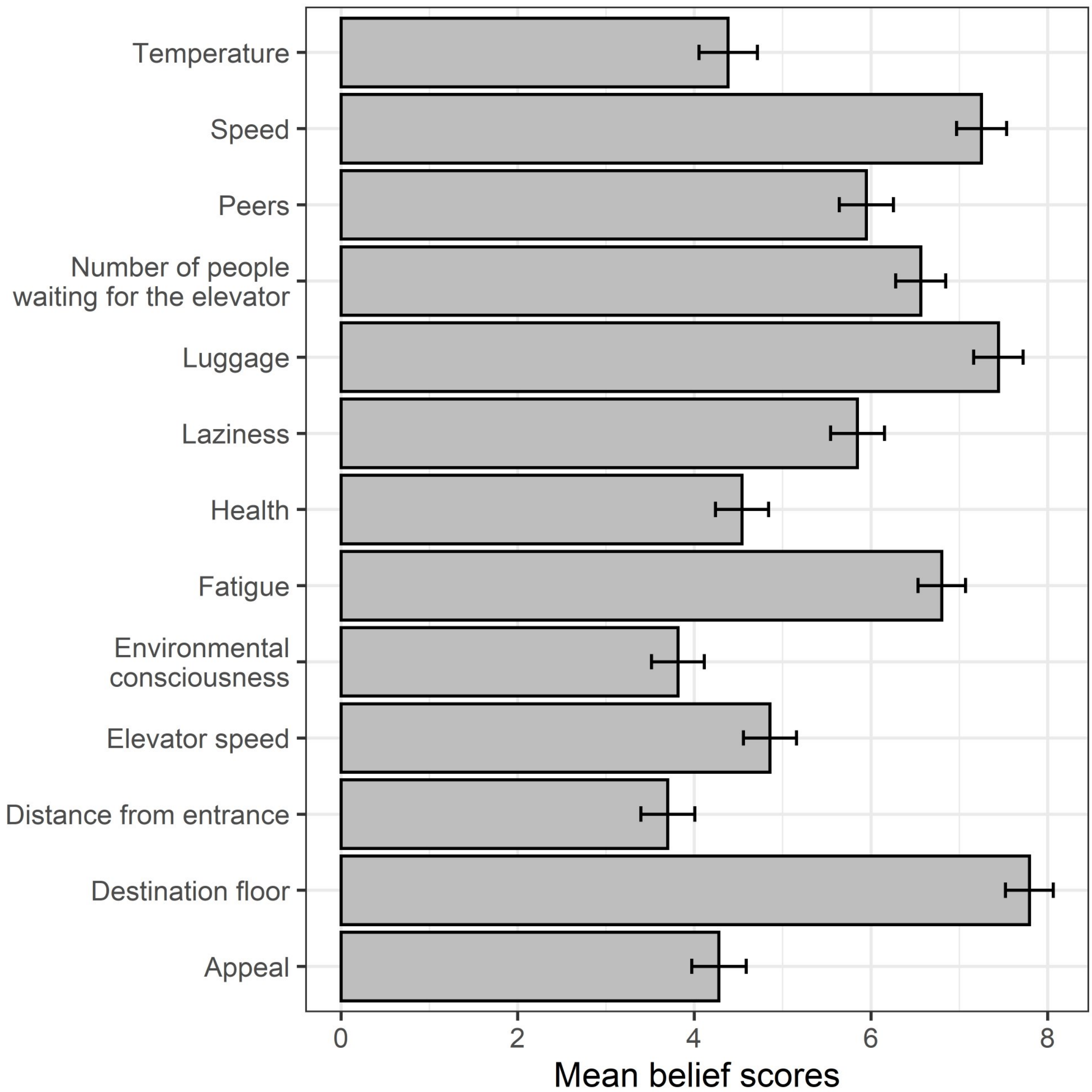

Figure 2. Believed importance mean scores of the potential influencing factors. Bars show the mean scores across subjects, while the error bars represent 95% confidence intervals.

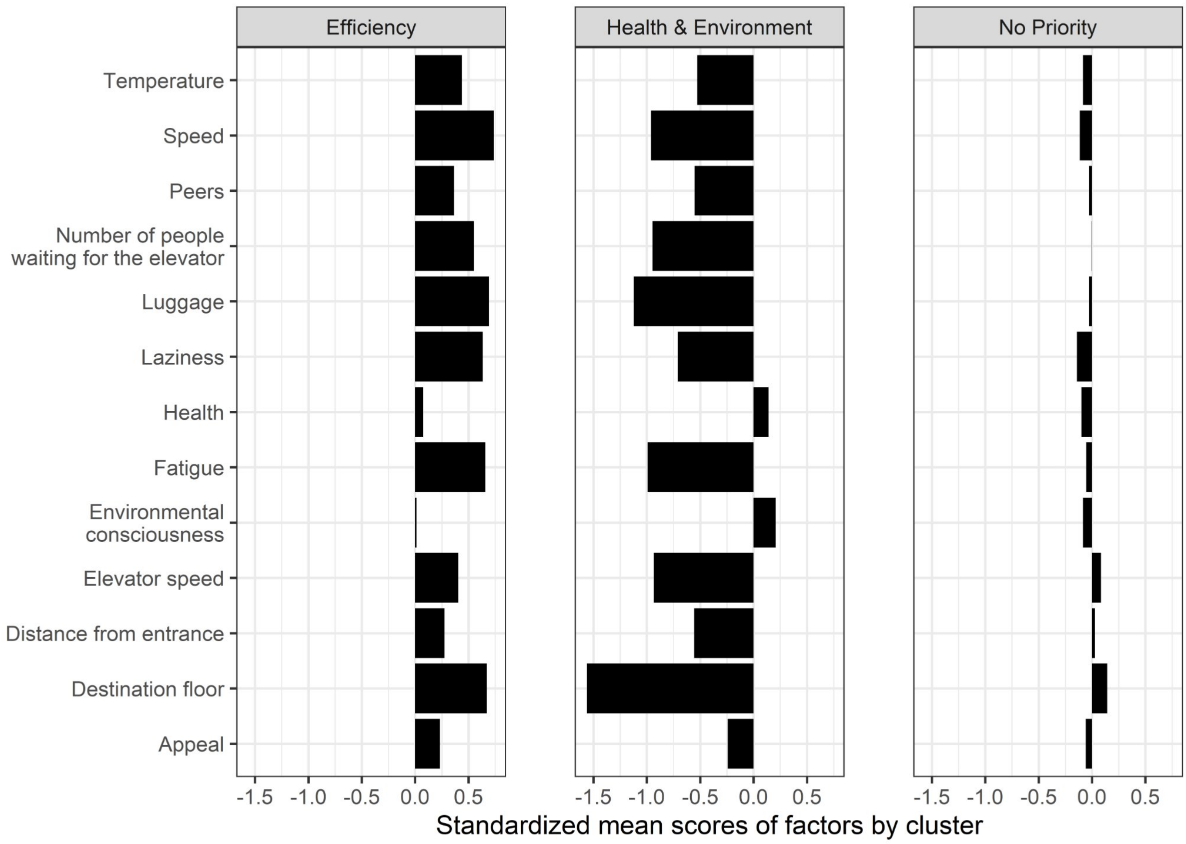

Figure 3. Believed importance standardized mean scores of the potential influencing factors by clusters. Bars show the standardized mean scores across subjects.

Discussion

The benefits of utilizing the Choice Context Exploration in this scenario are substantial. Without delving into the context of our decision, crucial influencing factors may have been overlooked, leaving intervention strategies without direction. Moreover, reliance solely on individuals' beliefs, potentially erroneous, could lead to misinterpretation of influential contextual factors. Behavioral measurements aid in identifying factors with the greatest impact, such as peer behavior, informing future intervention designs tailored to these contextual elements. This exploration method proves particularly advantageous in novel choice situations, environments, or populations, where traditional trial-and-error strategies may be impractical. By preemptively examining contextual influences, the expense and risk of ineffective interventions are minimized, enhancing the overall effectiveness of intervention efforts. Despite its merits, the Choice Context Exploration has limitations. It offers a general overview of contextual factors, potentially overlooking emerging influences or changes in factor effects over time. Additionally, reliance on self-reported data and the limited sample size pose constraints on generalizability. Further research is warranted to explore the applicability of this method across diverse environments and populations, considering moderating factors in nudge implementation.

Chapter 2

During pandemics when treatment or vaccines are unavailable, behavioral measures become crucial for containing diseases (Yang, 2021; Chen, 2021). One effective approach is to minimize contact by adhering to confinement recommendations, but sustaining compliance over time is challenging. Understanding the factors influencing compliance with these recommendations is vital for disease containment (Faulkner, 2021). This chapter focuses on contextual factors, defined as physical, sociocultural, and intrapersonal circumstances affecting decisions. These factors accurately predict decisions in simple scenarios. Identifying and understanding these factors regarding non-adherence to lockdown recommendations can inform interventions to mitigate risk. Lockdown regulations typically permit leaving residences for essential activities, while non-essential outings are considered non-compliant behavior. Mental states such as loneliness and boredom, along with beliefs about health precautions and trust in policies, influence compliance (Stickley et al., 2021; Boylan et al., 2021). Trust in regulations, social trust, and expert opinion also play roles. However, these factors have not been systematically studied across cultures. This research aims to fill this gap by investigating contextual factors influencing compliance across cultures. A pilot study identified potential factors, followed by a main study involving participants from 16 countries to assess how these factors predict compliance and risk-taking behaviors using a machine learning approach.

The main study aimed to assess how factors identified in the pilot study predict compliance with lockdown recommendations and the riskiness of out-of-home activities. Data collection and analysis methods were pre-registered and approved by the local institutional review board. Data were collected between April 29, 2020, and November 12, 2020, from 16 countries. Only data from countries with more than 100 respondents were analyzed (n = 42,283). Recruitment involved 16 research labs using various media outlets, university participant pools, or paid participant pools. Respondents reported demographic information and whether they left home for non-essential reasons in the previous 7 days. Participants rated statements on a Likert scale regarding factors identified in the pilot study. Event-specific items were presented only to those who left home during lockdown, assessing circumstances of their most recent outing. Event-general items were shown to all respondents. Attention-check items were included. The questionnaire was translated into 11 languages. Random forest models were used to predict non-compliance and activity riskiness. Models were created for each country separately. Non-compliance was predicted using event-general items, and riskiness was assessed based on event-specific and event-general items. Models were trained and tested on separate datasets, and variable importance was calculated to assess the influence of predictors. Root Mean Squared Errors were used to evaluate model accuracy.

After exclusions, the final sample for analysis included 42,283 individuals from 16 countries, with an average age of 40.92 years and 50.86% female. A heatmap (Figure 4) illustrates the relative importance of factors in predicting compliance with lockdown measures across countries. Feeling of responsibility was consistently among the top three most important factors in 13 out of 16 countries. Similarly, feelings of being caged at home and fear of getting infected were significant factors in 10 and 11 countries, respectively.

Figure 4. Variable importance values when predicting leaving home in each country.

Our models, based on event-general factors, accurately predicted whether individuals left their homes during lockdown. Prediction accuracy ranged from 62% in Portugal to 87% in Italy. Partial dependence plots (Fig 2) indicated that feeling of responsibility was negatively associated with non-adherence, while fear of infection decreased the likelihood of leaving home and feeling caged at home increased it.

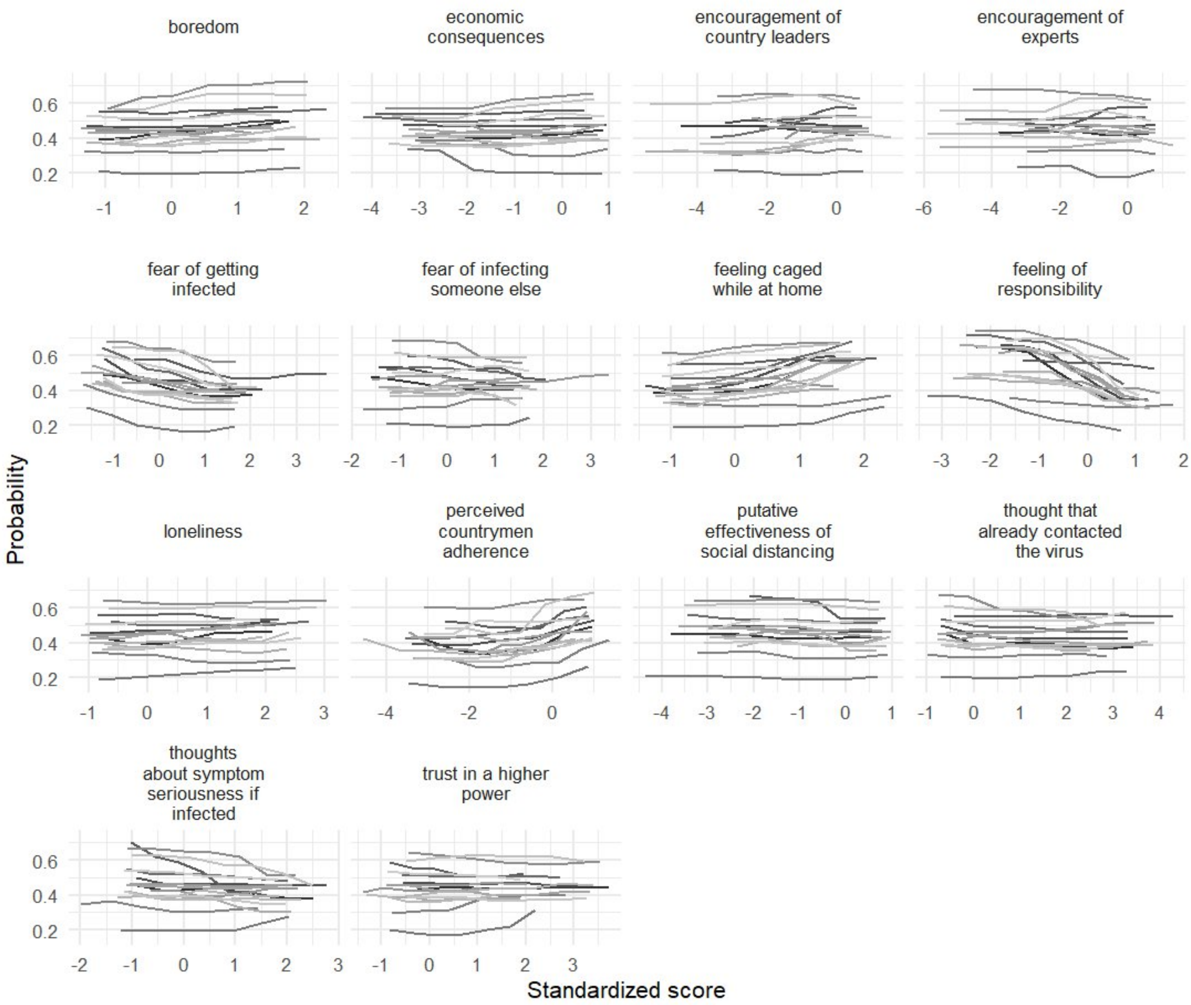

Figure 5. Partial dependence plots of variables used in the prediction of leaving home for all countries. Each line represents a different country.

Investigating factors associated with participation in risky activities, we found that the *putative effectiveness of social distancing, activity importance, and anticipated number of people met* were most important.

The study delved into the significance of contextual factors in predicting compliance with pandemic lockdown measures. Results revealed consistent patterns across 16 countries, indicating that certain factors either increased or decreased the likelihood of individuals leaving home during lockdowns. Notably, feelings of boredom and peer adherence to regulations consistently heightened the probability of leaving home, while fear of infection and a sense of responsibility decreased it across all countries. The prominence of the feeling of responsibility as a predictor of adherence underscores its importance in influencing behavior during crises. However, variations were observed among countries, with some factors proving more influential in specific contexts. For instance, while responsibility was a strong predictor in most countries, it mattered less in Japan, Switzerland, and Greece, where perceived adherence of fellow citizens held more sway. Similarly, fear of infection had varying effects across nations, suggesting cultural nuances in risk perception. The analysis of factors influencing activity riskiness further highlighted the importance of contextual considerations. Anticipated number of people met while traveling emerged as a key predictor, indicating that social interactions significantly contribute to activity risk. Trust in people met and activity importance also played notable roles, suggesting that individuals prioritize essential activities with trusted contacts, potentially mitigating risk.

Despite the study's comprehensive approach, certain limitations warrant consideration. Unidentified contributors to non-adherence were not accounted for, and context-specific factors may have been overlooked. Additionally, variations in sample sizes, data collection methods, and lockdown measures between countries could have influenced results. The inaccuracies in predicting activity riskiness underscore the complexity of this aspect and the need for further research. Overall, the findings underscore the importance of messaging that emphasizes personal responsibility and highlights the significance of individual actions in curbing disease transmission. Strategies to address social isolation and reframe confinement positively could enhance compliance. Transparent communication about symptoms and infection rates may also motivate adherence to regulations.

Chapter 3

In the face of an epidemic with no available medical treatment, persuading individuals to adopt behaviors that mitigate disease transmission becomes paramount (Betsch et al., 2020; Bavel et al., 2020). Social distancing, particularly the avoidance of social gatherings, has been recognized as an effective measure in curbing the spread of viruses, as evidenced by its recommendation or mandate in numerous countries during epidemics (Cheetham et al., 2021). However, the success of such mandates hinges on the speed and extent of their adoption within society. This study aims to identify demographic groups most likely to attend social gatherings during epidemic emergencies, thereby aiding public health officials and policymakers in designing targeted and effective campaigns.

Targeted interventions are essential for swift and efficient public health campaigns. Understanding key demographic groups enables policymakers to tailor interventions to specific contexts and characteristics, enhancing their effectiveness and potentially saving lives. In contrast, generic interventions may overlook population diversity, leading to decreased efficacy and avoidable negative outcomes. Furthermore, during pandemics, interventions targeting one group may inadvertently influence behavior in another, emphasizing the need for precise targeting to avoid unintended consequences.

While demographic factors such as age, gender, income, and education are readily available for targeting interventions, information on individuals' risk perceptions and norms is often lacking in traditional communication channels. However, these latent factors, including values and risk perceptions, play a crucial role in shaping social distancing behavior (Clark et al., 2020). Therefore, this study focuses on demographic factors readily accessible for targeting public health campaigns and policies.

We identify social gathering attendees based on age, education, gender, and income, as these factors have been shown to correlate with adherence behavior during epidemics. However, previous studies conducted in individual countries have yielded mixed results regarding the association between these variables and social distancing. The present research addresses this gap by investigating the association between demographic factors and social gathering avoidance in a large, diverse sample of over 80,000 individuals from 41 countries. Leveraging machine learning techniques, we not only identify main effects but also uncover subtle patterns and explore heterogeneity between countries. Ultimately, we discuss how these findings can inform and improve public health interventions.

The dataset used in this study was collected during the early phase of the COVID-19 pandemic, from March 20 to April 7, 2020, by an international research group. Using an online survey administered via the snowball method, participants were recruited from various social media and media outlets worldwide. A total of 112,136 individuals from 175 countries filled out the survey. To ensure data quality and relevance, several exclusion criteria were applied. Responses from individuals who did not complete the full survey were excluded. Additionally, responses were included only from countries where some form of restriction or recommendation affecting social gatherings was in effect at the time of the survey compilation. This information was verified using the Oxford COVID-19 Government Response Tracker, which collects publicly available data on COVID-19-related governmental responses in each country.

Further data cleaning involved removing responses with nonsensical values for age, household size, and years of education, as well as individuals reporting unrealistically high incomes. Finally, responses from countries with fewer than 400 participants were excluded to maximize the reliability of the dataset. The final dataset comprised 87,169 responses from 41 countries, with an average of 2,126 respondents per country. These countries represented 73.05% of the world's population as of 2020. Participants provided demographic information, including age, gender, education level, country of residence, monthly household income, and household size. Adjusted household income, calculated by dividing household income by the square root of household size, was used in the analyses. Participants responded to survey items related to COVID-19, including questions about their behavior regarding social gatherings over the past week, measured on a 100-point scale. Those indicating total agreement (100 points) with the statement "I did not attend social gatherings" were classified as social gathering avoiders, while the rest were classified as social gathering goers. Demographic information collected included age, gender, education level, country of residence, household income, and household size. Random forest models were employed to explore the role of demographics in social gathering avoidance. Random forests are robust to non-linearity and handle unbalanced data well compared to logistic regression models. Individual models were fitted to each country's data due to variations in disease progression, policy measures, and social norms. Data were split into training and test sets in an 80-20 ratio, with the number of variables sampled at each split of a decision tree tuned separately for each country via repeated 10-fold cross-validation. To address class imbalance, the training data were upsampled to balance social gathering avoiders and goers. Models were tuned on the training set to optimize the area under the precision-recall curve (prAUC), which is robust to unbalanced data. Model performance was evaluated using the test set.

Descriptive statistics were calculated to examine the proportion of social gathering goers and avoiders across different demographic subgroups in each country. This included subgroups based on gender, income level, age, and education. The analysis revealed that in 95% of the countries (39 out of 41), the proportion of social gathering goers was higher among males compared to females. In 80% of the countries (33 out of 41), social gathering goers were more prevalent among individuals with lower income compared to those with higher income. In 78% of the countries (32 out of 41), younger individuals were more likely to attend social gatherings compared to older individuals. In 66% of the countries (27 out of 41), individuals with lower education levels were more inclined to attend social gatherings compared to those with higher education levels.

Partial dependence plots were generated using the results from random forest models to visualize the association between each demographic factor and the probability of social gathering avoidance (Fig x). These plots showed that, in most countries, older age, being female, higher income, and more years of education were associated with a lower probability of attending social gatherings. However, there was significant heterogeneity across countries for each demographic variable except gender.

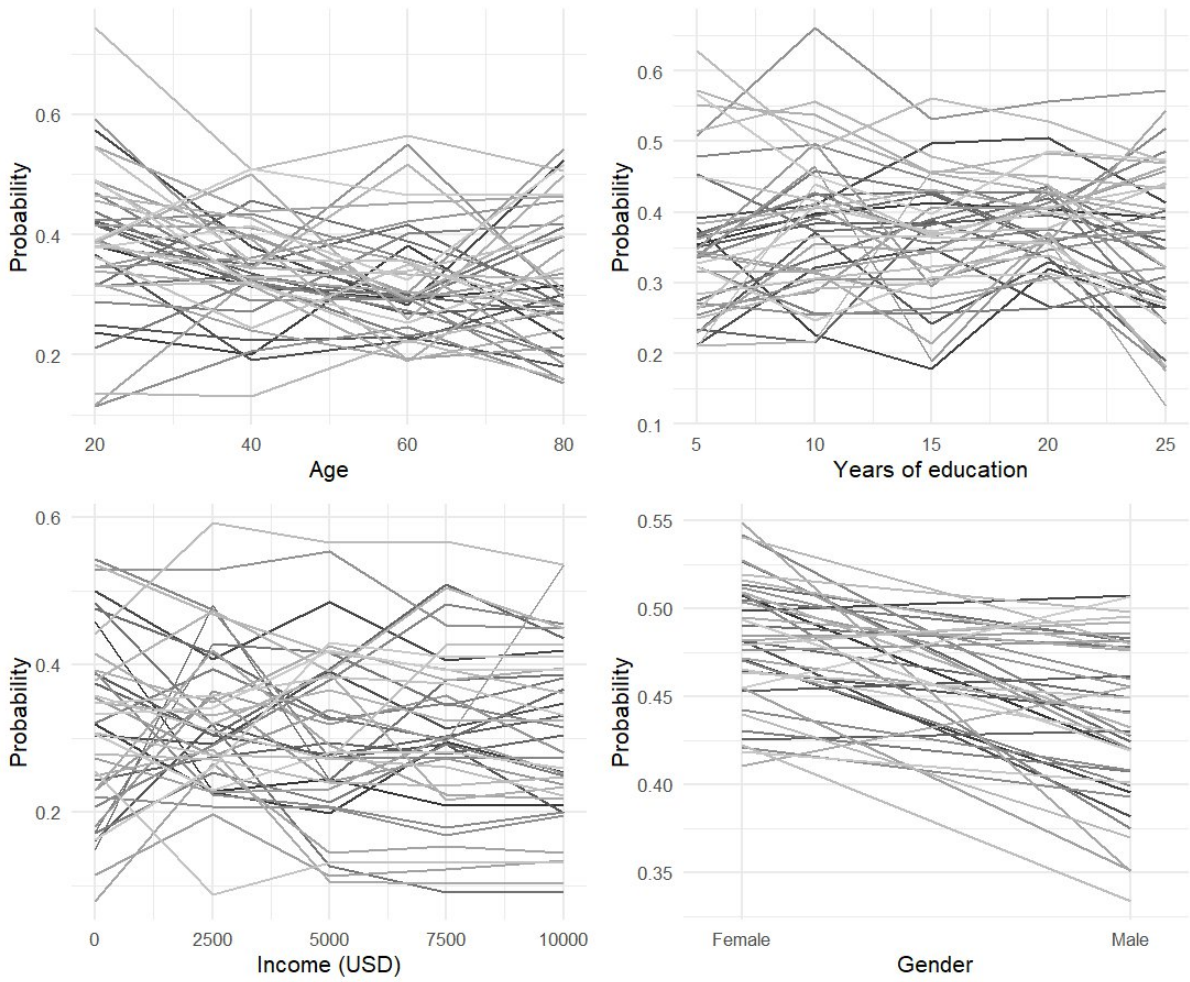

Figure 6. Partial dependence plots show the average predicted probability of leaving home associated with a given value of the demographic factor of age, years of education, income, and gender (in different plots) for all the countries.

Variable importance scores were calculated for each demographic factor in each country to assess their overall contribution to predicting social gathering avoidance. The median importance scores were as follows:

Income: Median importance score of 0.07 (range: 0 - 0.23) Age: Median importance score of 0.05 (range: 0 - 0.21) Education: Median importance score of 0.05 (range: 0 - 0.23) Gender: Median importance score of 0.02 (range: 0 - 0.13)

These scores indicate the percentage increase in prediction accuracy when each demographic factor is added to the model. The strongest predictor of social gathering avoidance was determined by identifying the demographic factor with the highest variable importance score in each country. Results showed that income was the strongest predictor in 29 out of 41 countries. Age was the strongest predictor in 10 countries. Education was the strongest predictor in 2 countries. Gender had the lowest importance score in 36 countries, followed by years of education in 4 countries, and income in 1 country.

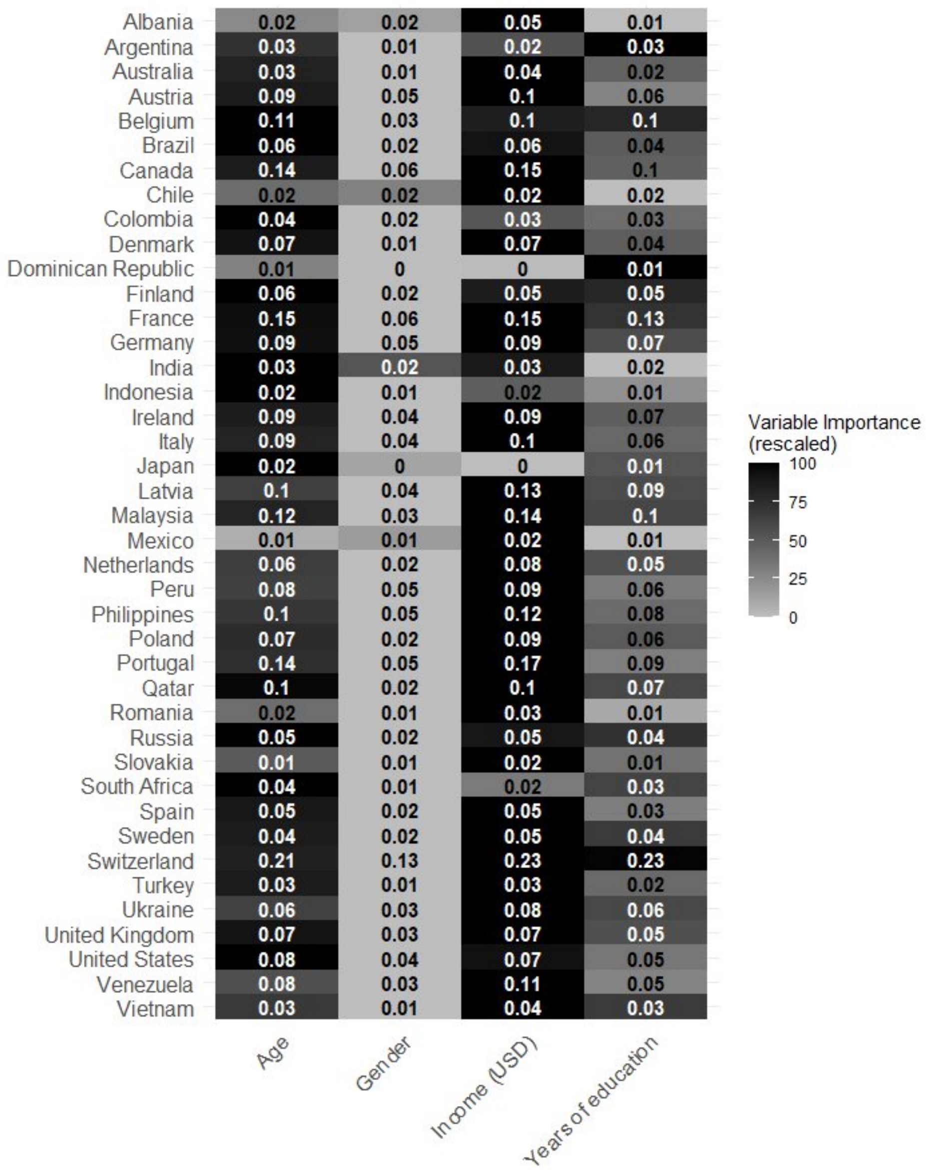

Figure 7. Variable importance scores for each demographic variable in each country. Variable importance values express the mean increase in accuracy when a given demographic variable is added to a model. The coloring of the figures depicts the relative importance of the variables within each country while the variable importance values were rescaled between 0 and 100 in each country, 100 being the most (darkest) and 0 being the least important (lightest).

Our study aimed to investigate the demographic factors influencing social gathering behavior during epidemic emergencies using a large dataset collected from 41 countries during the early phase of the COVID-19 pandemic. This global approach provides valuable insights into the diversity of social distancing behaviors across different demographic subgroups and countries, shedding light on potential targets for public health interventions. We found significant variations in social gathering behavior across demographic subgroups and countries. Generally, males, younger individuals, lower-educated individuals, and those with lower incomes were more likely to attend social gatherings. However, there was considerable heterogeneity among countries, indicating that generalized assumptions about social distancing behavior may not hold true globally. For instance, in some countries, higher-educated individuals exhibited higher rates of social gathering attendance compared to their lower-educated counterparts.

Income emerged as the strongest predictor of social gathering behavior across countries, followed by age, education, and gender. However, the relative importance of these factors varied widely among countries, highlighting the need for tailored interventions based on local contexts. Even neighboring countries exhibited different patterns, underscoring the importance of understanding country-specific dynamics.

Although the observed effects of demographic factors may seem small, they can have significant implications for disease transmission. For example, a small increase in social gathering attendance among certain demographic groups could lead to a substantial rise in new COVID-19 cases. Targeting these less adherent demographic subgroups with tailored public health campaigns could be a cost-effective strategy to mitigate transmission risks.

Our study has several limitations. Firstly, the data were collected during the early phase of the pandemic, and social distancing behavior may have evolved over time. However, evidence suggests that the associations between demographic factors and social distancing remained consistent over several months. Secondly, our reliance on self-reported data for social distancing behavior may introduce biases, although previous studies have shown that self-reports generally align with actual behavior. Thirdly, our findings may not generalize to future pandemics, as the factors driving social distancing decisions may vary between outbreaks. Finally, while we identified associations between demographic factors and social gathering behavior, further research is needed to elucidate the underlying mechanisms driving these patterns. In conclusion, our study underscores the importance of understanding the demographic drivers of social distancing behavior in epidemic emergencies. By identifying vulnerable demographic subgroups and tailoring interventions accordingly, policymakers and public health officials can effectively mitigate disease transmission and protect public health.

Chapter 4

The effectiveness of nudges often varies across different contexts, including populations, locations, cultures, and times. Despite this, there is a tendency to assume that results from experiments will generalize to other contexts, leading to misconceptions about the applicability of nudges. However, contextual factors play a crucial role in determining the success of behavioral interventions.

Each person has unique experiences, desires, and skills that influence their response to nudges. What resonates with one individual may not have the same effect on another. For example, a social norm message encouraging email usage may appeal to some but be ineffective for others due to differences in perceptions or familiarity with technology. Even the same individual may respond differently to nudges based on the situation or time of day. Factors such as stress levels or daily routines can impact receptiveness to behavioral cues. Understanding these situational nuances is essential for designing effective interventions. The effectiveness of nudges can be illustrated by real-world examples, such as attempts to reduce energy consumption through social norms. While initial studies showed promising results, scaling up interventions revealed significant contextual differences that diminished effectiveness. Factors like household demographics and attitudes played a crucial role in shaping behavior. In our own research, we found that contextual factors influenced people's choices between taking the stairs or the elevator. This underscores the importance of considering context when designing and implementing behavioral interventions. Contextual factors likely play a similar role in influencing the effectiveness of nudges.

Understanding context is critical for distinguishing between successful and failed nudges. Recent studies have shown that the effectiveness of nudges can vary widely, with some interventions having minimal impact despite initial expectations. Acknowledging failures and learning from them is essential for improving the effectiveness of behavioral interventions. Recognizing the importance of contextual variability is essential for successful implementation of nudges. By understanding individual differences, situational nuances, and real-world examples, practitioners can design interventions that are more likely to achieve the desired behavioral outcomes. Learning from both successes and failures is key to advancing the field of behavioral science and maximizing the impact of nudges.

Understanding the context in which a behavioral intervention will be implemented is crucial for its success. Contextual factors can include physical attributes of the environment, social and cultural norms, psychological attributes and preferences of the target population, timing of the decision or behavior, and characteristics of the intervention itself. Conducting a structured exploration of these factors is essential before implementing a behavioral intervention. Drawing inspiration from human-centered design, organizations can explore and understand their target groups more effectively. By analyzing behavioral data and conducting qualitative research such as interviews or focus groups, decision-makers can identify key contextual factors that may influence the effectiveness of the intervention.

Once contextual factors have been identified, it is essential to test the intervention with diversity in mind. Randomized testing should be conducted on a small sample in a context similar to the target population. However, it's not just about testing the intervention's effectiveness overall; it's also about understanding how its effects may vary across different subgroups within the population.

We should anticipate how the intervention's effectiveness might vary across different subgroups of the population. It is considered good practice to divide the population into subgroups based on relevant contextual factors and recruit participants from each subgroup for testing. This allows for a more nuanced understanding of how the intervention performs across different demographic or behavioral segments.

We also have to ensure that each subgroup has a large enough sample size to estimate the average effect accurately. By testing the intervention in each subgroup, decision-makers can identify which segments of the population are most responsive to the nudge and tailor implementation strategies accordingly. By carefully considering contextual factors and testing

with diversity in mind, decision-makers can increase the likelihood of success for their behavioral interventions.
Unlike in a lab where the context can be controlled, when applying behavioural insights in the wild, we constantly run into new configurations of the factors that might have an impact both on a target behaviour and on the effectiveness of a nudge. Assuming that the effectiveness of the behavioral intervention will vary and may even not work is possibly the best motivation for any behavioural scientist to keep exploring the context of a behaviour. We should not get carried away with any testing opportunity but instead focus on systematically building a diverse test sample that ensures we will end up being able to tell successful and failed nudges apart. After all, when exploring the wild, preparing for surprises is the best strategy one can follow.

Discussion

What have we achieved by incorporating machine learning techniques and algorithms into this research? The practice of splitting our data into training and test sets has proven to be effective in mitigating overfitting, as demonstrated in Chapter 1. Splitting data into training and test sets is a common practice in machine learning and statistical modeling. It involves dividing a dataset into two distinct subsets for the purpose of developing and evaluating a predictive model. The training set is a portion of the dataset that is used to train and build the predictive model. It contains a significant portion of the data and serves as the foundation for the model to learn patterns, relationships, and trends within the data. The model is fitted to this training set, which means it adjusts its parameters and algorithms to capture the underlying characteristics of the data. The test set is a separate portion of the dataset that is held back and not used during the training phase. Instead, it is used to assess the performance of the model after it has been trained. The test set is used to simulate the real-world application of the model, allowing you to evaluate how well the model generalizes from the training data to make predictions on new, unseen data.

The primary purpose of splitting data into training and test sets is to evaluate the model's ability to make accurate predictions on data it has not been exposed to during training. It helps identify whether the model has learned to recognize genuine patterns in the data or if it has overfit the training data (meaning it has memorized the training data but cannot generalize to new data). This process is critical for assessing the model's performance, making adjustments if necessary, and ensuring that it is reliable for making predictions in real-world scenarios.

The accuracy of a model, regardless of the specific accuracy metric employed, consistently appears higher on a training set than on a test set. By partitioning our data in this manner, we have prevented an overestimation of accuracy. However, it is worth noting that in Chapter 3, the random forest algorithm did not surpass the performance of the logistic regression models. This suggests that non-linear effects, which the random forest algorithm is designed to detect but the logistic regression algorithm may miss, were not present in our dataset. The value of the random forest algorithm becomes particularly evident when dealing with datasets that exhibit non-linear relationships between variables and a substantial number of predictors. In our specific case, neither of these conditions applied. Random forests can, at times, exhibit overfitting, as observed in certain models in Chapter 2. Nevertheless, the instances where models clearly overfitted, as evidenced by their lower accuracy compared to the base rate, were relatively few. All in all, our research indicates that machine learning tools can be useful and informative when used during exploratory analyses. However, these methods are not extensively used in the analysis of psychological experiments as compared to other fields, for example genetics (Orrù et al., 2020). What are the possible reasons behind this?

There are concerns regarding the interpretability and transparency of machine learning models in psychology research. Traditional statistical methods often produce results that are easier to interpret and explain, whereas machine learning models, particularly complex ones like neural networks, are often perceived as "black boxes." This lack of transparency can be problematic in fields like psychology, where understanding the underlying mechanisms of behavior is crucial. Researchers may be hesitant to adopt machine learning techniques if they cannot fully understand or explain the reasoning behind the model's predictions. Ethical and privacy concerns also play a role in limiting the widespread use of machine learning in psychology research. Machine learning models rely on large amounts of data, raising questions about data privacy, security, and potential biases in the data. Researchers must navigate these ethical considerations carefully, especially when working with sensitive psychological data. Furthermore, cultural and institutional factors within the field of psychology may contribute to the slow adoption of machine learning techniques. Resistance to change, disciplinary boundaries, and academic incentives may all influence research practices and hinder interdisciplinary collaboration.

Overall, while machine learning holds great promise for advancing our understanding of human behavior, its integration into psychology research faces several challenges. Addressing these barriers will require efforts to increase awareness and understanding of machine learning techniques, provide training and education opportunities, promote interdisciplinary collaboration, address ethical and privacy concerns, and overcome cultural and institutional barriers within the field. By addressing these challenges, psychologists can harness the power of machine learning to gain deeper insights into human behavior and improve the effectiveness of psychological interventions.

Further directions

To further leverage the power of machine learning in this domain, several avenues for future research can be explored. Firstly, prioritizing the collection of diverse and comprehensive datasets is essential. By expanding the range of variables and incorporating multimodal data sources, researchers can gain a more holistic understanding of human behavior. Feature engineering techniques, such as feature selection and extraction, can further enhance the informativeness of variables, revealing latent patterns within the data.

Secondly, employing different analyses, such as contextual analysis, natural language processing (NLP), temporal analysis, and longitudinal studies (Evans et al., 2023), can offer valuable insights into decision-making processes and behavior dynamics over time. Integrating multimodal data and leveraging advanced machine learning methods can provide a more comprehensive view of human behavior and its underlying mechanisms (Akkus et al., 2023).

Thirdly, semi-supervised and active learning techniques can be valuable in scenarios where labeled data is limited or expensive to obtain. These approaches enable models to learn from both labeled and unlabeled data, reducing the annotation burden and optimizing resource utilization. Furthermore, reinforcement learning can be applied to optimize interventions and treatments over time in exploratory studies involving sequential decision-making processes. Scalable machine learning platforms offer the necessary computational infrastructure to handle large datasets and complex models effectively.

Finally, establishing causation, rather than merely identifying correlations, remains a fundamental challenge in exploratory research. Counterfactual analysis, causal inference techniques, and randomized controlled trials (RCTs) done with the help of machine learning can facilitate the identification of causal relationships and deepen our understanding of the factors influencing human behavior. Overall, by embracing interdisciplinary collaboration, leveraging advanced analytical techniques, and addressing ethical and methodological challenges, researchers can unlock new insights into the complex dynamics of human behavior and pave the way for more effective interventions and treatments in psychology and related fields.

Within the context of utilizing machine learning in exploratory psychology research, one emerging area of paramount importance is *Explainable AI* (XAI; Arrieta et al., 2020). When employing machine learning models, especially deep learning techniques, researchers often encounter the "black box" problem, where predictions lack comprehensible rationale. This opacity obstructs understanding the decision-making mechanisms, hindering hypothesis generation and bias detection. XAI addresses these challenges by rendering AI systems more transparent and comprehensible. Techniques like Local Interpretable Model-agnostic Explanations (LIME) generate human-readable explanations, allowing researchers to identify which features influenced specific predictions. This transparency facilitates bias detection, enabling proactive mitigation and ensuring fairness in research outcomes.

Ethical considerations are paramount as machine learning's role expands in exploratory research. Prioritizing ethical data collection, ensuring algorithm transparency, and implementing bias mitigation techniques are essential to address concerns related to data privacy, bias, and fairness. Fairness-aware machine learning and open science practices promote reproducibility and inclusivity while safeguarding against unethical use of research assets. Collaboration between machine learning experts and domain-specific researchers fosters synergistic interactions, amplifying the impact of studies. Human-AI collaboration platforms empower researchers to leverage AI systems in hypothesis generation, pattern identification, and task automation, freeing experts to focus on interpretation and theory building (Wang et al., 2020).

In conclusion, embracing complexity in behavioral science unlocks new avenues for understanding human behavior and developing effective interventions. By navigating these diverse directions, researchers contribute to a deeper understanding of human decision-making and behavior, advancing knowledge boundaries in these domains. This evolution promises transformative impacts on societal challenges and research frontiers alike.

References

Agrawal, M., Peterson, J. C., & Griffiths, T. L. (2020). Scaling up psychology via scientific regret minimization. *Proceedings of the National Academy of Sciences*, *117*(16), 8825–8835.

Akkus, C., Chu, L., Djakovic, V., Jauch-Walser, S., Koch, P., Loss, G., Marquardt, C., Moldovan, M., Sauter, N., Schneider, M., Schulte, R., Urbanczyk, K., Goschenhofer, J., Heumann, C., Hvingelby, R., Schalk, D., & Aßenmacher, M. (2023). *Multimodal Deep Learning*. https://doi.org/10.48550/ARXIV.2301.04856

Arrieta, A., Díaz-Rodríguez, N., Del Ser, J., Bennetot, A., Tabik, S., Barbado, A., Garcia, S., Gil-Lopez, S., Molina, D., Benjamins, R., Chatila, R., & Herrera, F. (2020). Explainable Artificial Intelligence (XAI): Concepts, taxonomies, opportunities and challenges toward responsible AI. *Information Fusion*, *58*, 82–115. https://doi.org/10.1016/j.inffus.2019.12.012

Bavel, J. J. V., Baicker, K., Boggio, P. S., Capraro, V., Cichocka, A., Cikara, M., ... & Willer, R. (2020). Using social and behavioural science to support COVID-19 pandemic response. *Nature human behaviour*, *4*(5), 460-471.

Betsch, C., Korn, L., Sprengholz, P., Felgendreff, L., Eitze, S., Schmid, P., & Böhm, R. (2020). Social and behavioral consequences of mask policies during the COVID-19 pandemic. *Proceedings of the National Academy of Sciences*, *117*(36), 21851-21853.

Borsboom, D. (2017). A network theory of mental disorders. *World Psychiatry*, *16*(1), 5–13.

Borsboom, D., van der Maas, H. L., Dalege, J., Kievit, R. A., & Haig, B. D. (2021). Theory construction methodology: A practical framework for building theories in psychology. *Perspectives on Psychological Science*, *16*(4), 756–766.

Boylan, J., Seli, P., Scholer, A. A., & Danckert, J. (2021). Boredom in the COVID-19 pandemic: Trait boredom proneness, the desire to act, and rule-breaking. *Personality and individual differences*, *171*, 110387.

Brandon, A., Ferraro, P., List, J., Metcalfe, R., Price, M., & Rundhammer, F. (2017). *Do The Effects of Social Nudges Persist? Theory and Evidence from 38 Natural Field Experiments* (No. w23277; p. w23277). National Bureau of Economic Research. https://doi.org/10.3386/w23277

Bryan, C. J., Tipton, E., & Yeager, D. S. (2021). Behavioural science is unlikely to change the world without a heterogeneity revolution. *Nature Human Behaviour*, *5*(8), 980–989. https://doi.org/10.1038/s41562-021-01143-3

Cheetham, N., Waites, W., Ebyarimpa, I., Leber, W., Brennan, K., & Panovska-Griffiths, J. (2021). Determining the level of social distancing necessary to avoid future COVID-19 epidemic waves: a modelling study for North East London. *Scientific reports*, *11*(1), 5806.

Chen, R. M. (2021). Whether County Lockdown Could Deter the Contagion of COVID-19 in the USA. *Risk Management and Healthcare Policy*, 2665-2673.

Clark, C., Davila, A., Regis, M., & Kraus, S. (2020). Predictors of COVID-19 voluntary compliance behaviors: An international investigation. *Global transitions*, *2*, 76-82.

Cramer, A. O., Van Borkulo, C. D., Giltay, E. J., Van Der Maas, H. L., Kendler, K. S., Scheffer, M., & Borsboom, D. (2016). Major depression as a complex dynamic system. *PloS One*, *11*(12), e0167490.

De Groot, A. D., & Spiekerman, J. A. (2020). *Methodology: Foundations of inference and research in the behavioral sciences* (Vol. 6). Walter de Gruyter GmbH & Co KG.

Duflo, E., Kremer, M., & Robinson, J. (2011). Nudging Farmers to Use Fertilizer: Theory and Experimental Evidence from Kenya. *American Economic Review*, *101*(6), 2350–2390. https://doi.org/10.1257/aer.101.6.2350

Evans, S. L., Jones, R., Alkan, E., Sichman, J. S., Haque, A., de Oliveira, F. B. S., & Mougouei, D. (2023). The emotional impact of COVID-19 news reporting: A longitudinal study using natural language processing. *Human Behavior and Emerging Technologies*, *2023*.

Green, C. D. (2021). Perhaps Psychology's Replication Crisis is a Theoretical Crisis that is Only Masquerading as a Statistical One. *International Review of Theoretical Psychologies*, *1*(2), Article 2. https://doi.org/10.7146/irtp.v1i2.127764

Faulkner P. Lockdown: a case study in how to lose trust and undermine compliance. Global Discourse. 2021 May 1;11(3):497–515.

Hajdu, N., Szaszi, B., Aczel, B., & Nagy, T. (2023, September 18). Using supervised machine learning methods in psychological research. https://doi.org/10.31234/osf.io/tjkug

John, P., Cotterill, S., Richardson, L., Moseley, A., Smith, G., Stoker, G., Wales, C., Liu, H., & Nomura, H. (2013). *Nudge, nudge, think, think: Experimenting with ways to change civic behaviour*. A&C Black.

Jordan, M. I., & Mitchell, T. M. (2015). Machine learning: Trends, perspectives, and prospects. *Science*, *349*(6245), 255–260.

Kruglanski, A. W. (2001). That" vision thing": The state of theory in social and personality psychology at the edge of the new millennium. *Journal of Personality and Social Psychology*, *80*(6), 871.

Mischel, W. (2008). The Toothbrush Problem. *APS Observer*, *21*. https://www.psychologicalscience.org/observer/the-toothbrush-problem

Nosek, B. A., Alter, G., Banks, G. C., Borsboom, D., Bowman, S. D., Breckler, S. J., Buck, S., Chambers, C. D., Chin, G., Christensen, G., Contestabile, M., Dafoe, A., Eich, E., Freese, J., Glennerster, R., Goroff, D., Green, D. P., Hesse, B., Humphreys, M., … Yarkoni, T. (2015). Promoting an open research culture. *Science*, *348*(6242), 1422–1425. https://doi.org/10.1126/science.aab2374

Rogers, L., De Brún, A., & McAuliffe, E. (2020). Defining and assessing context in healthcare implementation studies: a systematic review. *BMC Health Services Research*, *20*(1), 1-24.

Scheel, A. M. (2022). Why most psychological research findings are not even wrong. *Infant and Child Development*, *31*(1), e2295.

Scheel, A. M., Schijen, M. R. M. J., & Lakens, D. (2021). An Excess of Positive Results: Comparing the Standard Psychology Literature With Registered Reports. *Advances in Methods and Practices in Psychological Science*, *4*(2), 251524592110074. https://doi.org/10.1177/25152459211007467

Silva, A., & John, P. (2017). Social norms don't always work: An experiment to encourage more efficient fees collection for students. *PLOS ONE*, *12*(5), e0177354. https://doi.org/10.1371/journal.pone.0177354

Stickley, A., Matsubayashi, T., & Ueda, M. (2021). Loneliness and COVID-19 preventive behaviours among Japanese adults. *Journal of Public Health*, *43*(1), 53-60.

Szaszi, B., Palinkas, A., Palfi, B., Szollosi, A., & Aczel, B. (2018). A Systematic Scoping Review of the Choice Architecture Movement: Toward Understanding When and Why Nudges Work: Systematic Scoping Review of the Nudge Movement. *Journal of Behavioral Decision Making*, *31*(3), 355–366. https://doi.org/10.1002/bdm.2035

Szollosi, A., & Donkin, C. (2021). Arrested theory development: The misguided distinction between exploratory and confirmatory research. *Perspectives on Psychological Science*, *16*(4), 717–724.

Vallacher, R. R., & Nowak, A. (1997). The emergence of dynamical social psychology. *Psychological Inquiry*, *8*(2), 73–99.

Van Lissa, C. J. (2022). Developmental data science: How machine learning can advance theory formation in Developmental Psychology. *Infant and Child Development*, e2370.

Wang, D., Churchill, E., Maes, P., Fan, X., Shneiderman, B., Shi, Y., & Wang, Q. (2020, April). From human-human collaboration to Human-AI collaboration: Designing AI systems that can work together with people. In *Extended abstracts of the 2020 CHI conference on human factors in computing systems* (pp. 1-6).

Wicherts, J. M., Veldkamp, C. L. S., Augusteijn, H. E. M., Bakker, M., van Aert, R. C. M., & van Assen, M. A. L. M. (2016). Degrees of Freedom in Planning, Running, Analyzing, and Reporting Psychological Studies: A Checklist to Avoid p-Hacking. *Frontiers in Psychology*, *7*. https://www.frontiersin.org/articles/10.3389/fpsyg.2016.01832

Yarkoni, T., & Westfall, J. (2017). Choosing Prediction Over Explanation in Psychology: Lessons From Machine Learning. *Perspectives on Psychological Science*, *12*(6), 1100–1122. https://doi.org/10.1177/1745691617693393

Yang, X. (2021). Does city lockdown prevent the spread of COVID-19? New evidence from the synthetic control method. *Global Health Research and Policy*, *6*, 1-14.