INSIGHTS ON IMMIGRATION, PERSONALITY AND LANGUAGE USE

IMMIGRATION, PERSONALITY AND LANGUAGE USE:

INSIGHTS FROM THREE LARGE-SCALE ANALYSES

By DAVIDE GENTILE, B.A.

A Thesis Submitted to the School of Graduate Studies in Partial Fulfillment of the Requirements

for the Degree Master of Science

McMaster University MASTER OF SCIENCE (2019) Hamilton, Ontario (Cognitive Science of Language)

TITLE: Immigration, Personality and Language Use: Insights from three large-scale Analyses

AUTHOR: Davide Gentile, B.A. (University of Bologna)

SUPERVISOR: Dr. V. Kuperman

NUMBER OF PAGES: ix, 75

Lay Abstract

The goal of this thesis is to gain insight on the psychology of migratory populations. We conducted three large scale empirical studies to address different questions about immigration, personality and language use. Immigrants and non-immigrants show differences throughout some dimensions of the Five Factor model of personality. We observed which personality traits can predict different migration distances, but also how one's migration affects their levels of openness. Finally, we examined whether differences in personality of immigrants and non-immigrants are reflected in their language use on social media. Taken together, our results contribute with both new evidence and insights to the knowledge about the psychological components of immigration and differences in language use among immigrants and non-immigrants.

Abstract

The goal of this thesis is to gain insight on the psychology of migratory populations. Immigrants and non-immigrants show differences throughout some dimensions of the Five Factor model of personality. We conducted three large scale empirical studies to address different questions about immigration, personality and language use. In Chapter 1, we use personality scores and geographical information from more than 1.5 million responders to investigate which traits can predict different migration distances within and across national borders. Our results highlight the mechanisms of how one's transitory psychological configuration can assist, accompany or interfere with the act of migration. In Chapter 2, we examine how one's migration affects their levels of openness, and how this relationship is modulated by other known determinants of personality traits such as age and gender. Our findings shed light on the psychological effects of immigration and consider how these effects are modulated by people of different age and gender. Finally, in Chapter 3 we use demographic, psychological and textual data from Facebook to determine how differences in personality of immigrants and non-immigrants are reflected in their language use on Facebook. The comprehensive exploration carried out in Chapter 3 gives insights on how language use distinguishes people of different immigration status. Taken together, our results contribute with both new evidence and insights to the knowledge about the psychological components of immigration and differences in language use among immigrants and non-immigrants.

Table of Contents

List of Figures

List of Tables

Declaration of Academic Achievement

The conceptualization of this thesis was conducted by myself, in collaboration with Dr.

Kuperman. I explored and analyzed the data under the supervision of Dr. Kuperman. The

statistical analysis was completed by myself and Dr. Kuperman. Contributions to the each of the

articles we intend to submit will be noted in the preface to those chapters. The present thesis was

written by myself, with comments from Dr. Kuperman.

0 Introduction

The goal of this thesis is to gain insight on the psychology of migratory populations. This thesis is comprised of three independent studies, each with specific backgrounds, data, questions and methods. The present introductory section provides an overview of the general concepts addressed in the body of the thesis. Details related to the specific studies are discussed independently in each chapter. The thesis is structured as follows. First, we introduce the main topic and the issues related. Then, we present the three studies separately, discussing background, methods and results for each chapter. Finally, we sum up our findings and draw general conclusions in the final discussion section.

The relationship between immigration and personality

Immigrants and immigration are subjects of study in many disciplines, from anthropology (Brettell, 2013; Brettell & Hollifield, 2013) to economics (Bodvarsson & Van den Berg, 2013; Borjas, 2014) to political studies (Freeman, 1995; Hollifield, 2000). In this scenario, investigating the psychology of migratory populations helps understanding the motivations and possibly the aspirations of people who experience migration. The subfield in psychology that examines the behaviour of migratory populations is usually labelled either 'migration psychology' or 'psychology of immigration' (Berry, 2001; Fawcett, 1985).

Migrations differ in nature, i.e., they can be permanent or temporary and occur within the same city or from one country to another. Similarly, the reasons behind one's migration are subject to a lot of variability (Boyle, 2009; Stark & Taylor, 1989). Generally, the causes of international migration are investigated in terms of economic causes, such as differentials in

work demand and supply between countries. However, this leaves uncovered the motivations of those who, prior the migration, were living in relatively wealthy country. To better explain individual differences in international migration, Boneva and Frieze (2001) introduced the idea of 'migrant personality'. According to their view, while economic factors might create the conditions for wanting to leave, the desire to do so is determined by individual choices and personality.

Previous studies have linked differences in personality to a number of aspects regarding migration, such as the distance covered by one's migration (Jokela et al., 2008; Jokela, 2009; Murray et al., 2005) and differences in migration intention (Camperio Ciani et al., 2007; Silventoinen et al., 2007; Tabor, 2010; Tabor & Milfont, 2011). In such studies, personality is commonly assessed through the adoption of the Five Factor Model of personality traits. The Five Factor Model follows the idea that one's temperament and behavior can be estimated in terms of individual scores in each of the Big Five personality traits: openness, conscientiousness, extraversion, agreeableness and neuroticism (Barrick & Mount, 1991; Goldberg, 1990). The present thesis focuses on three separate aspects of migration and personality: the relationship between personality traits and distance of migration, the effect of migration on openness and finally differences in language use between immigrants and non-immigrants on social media.

Content outline

Chapter 1 investigates the relationship between the Big Five traits and different distances of migration. We used a large-scale data set with Big Five personality scores and geographical information from millions of responders and examined whether the prominence of certain traits

could predict how far people went when they migrated. As anticipated, previous studies showed that personality traits can predict whether individuals migrate within or across countries. The original contribution in chapter 1 is in the inclusion of a previously unstudied international type of migration (iv) and a direct comparison of personality scores across all migration types (from non-migration to national and international migration).

The questions we address in the next two chapters lead us to constrain our focus on the only Big Five trait which seems to be consistently associated with migration propensity: *openness to experience.* In fact, several studies reported openness to experience as a valid predictor of not only migration propensity (Camperio Ciani et al., 2007; Canache et al., 2013; Gentile et al., 2019; Jokela 2008, 2009; Silventoinen et al., 2008; Tabor, 2015), but also foreign language learning achievement (Ozanska & Dewaele, 2012, Verhoeven & Vermeer 2002) and willingness to relocate for a job among unemployed individuals (Otto & Dalbert, 2012). These findings suggest that immigrants' openness is a necessary condition for mobility, and therefore relevant to individual differences in decision-making processes, migration and successful integration.

In chapter 2, we examine how immigration status of an individual (determined based on whether they are immigrants or not) affects their levels of openness, and how this relationship is modulated by other known determinants of personality traits: age and gender. The notion that personality changes over the course of adulthood due to adaptation to new environments is commonly accepted in the literature (e.g., Haan et al., 1986; Helson, Mitchell, & Moane, 1984; Neugarten, 1972). In chapter 2 we propose and argue that migration might be one of the factors modulating people's changes in personality over a lifetime. According to this view, one's levels of openness would differ based on whether that individual is an immigrant or not. In the first

analysis we observe whether immigration has an effect on openness levels. Then, given that differences in personality traits are found between different age groups (Roberts & Mroczek, 2008; Soto, John, Gosling & Potter, 2011) and genders (Helson, Pals, & Solomon, 1997; Weisberg, DeYoung, & Hirsh, 2011; Wink & Helson, 1993), we also ask how the influence of immigration is modulated by these two major contextual variables.

Finally, in chapter 3 we used demographic, psychological and textual data from Facebook to determine whether differences in personality of immigrants and non-immigrants are reflected in their language use on Facebook. Language and psychology are strictly connected; in fact, observing trends and differences in language use represents an additional tool to explore the psychology and behaviour of people (Coltheart, 1981; Pennebaker et al., 2003; Stone et al., 1996). Previous studies explored the language of social media to identify words and topics that classify groups of people based on variables of interest, such as geographical location, gender, age and psychological characteristics (Snefjella et al., 2018; Schwartz et al., 2013). By finding words and topics that distinguish immigrants and non-immigrants, with Chapter 3 we will shed additional light on the linguistic behaviour and psychology of both immigrants and non-immigrants.

Although we use the same dataset for chapters 1 and 2, we apply different restrictions to the data, resulting in different sample sizes and variables of interest. Moreover, apart from rare exceptions, literature reviews for chapter 1 and 2 do not overlap: as mentioned, data and questions are independent from each other. To conclude, there is no repetition of background or methodology for each of three chapter.

1.0 Big Five traits as predictors of migration distance in Canada and the US

Davide Gentile[1], Constance Imbault[1], Samuel Gosling[2],

Jason Rentfrow[3], Jeff Potter[4] & Victor Kuperman[1]

[1]McMaster University

[2]University of Texas at Austin

[3]University of Cambridge

[4]Enveritas

1.1 Preface

This article has been prepared to be submitted as a brief report to the Journal of Research in Personality. The data were collected by Dr. Samuel Gosling, Dr. Jason Rentfrow and Dr. Jeff Potter, who agreed to share them with us. The conceptualization of this report was conducted by myself, in collaboration with Dr. Kuperman. I explored, cleaned and analyzed the data under the supervision of Dr. Kuperman. The statistical analysis was completed by myself and Dr. Kuperman. The article was written by myself, with guidance and edits from Dr. Victor Kuperman and feedback from Constance Imbault and Dr. Jason Rentfrow.

1.2 Abstract

Personality traits can predict both the propensity of people to migrate and how far they migrate within and across countries. We used a large-scale data set with Big Five personality scores and geographical information from millions of responders to examine psychology of migratory populations. Our study (N = 1,674,852) provided new evidence that different distances of migration are associated with different personality traits. We identified four groups of populations currently residing in the US or Canada: individuals who (i) never changed residency, (ii) moved within a state or province, (iii) moved across states or provinces or (iv) moved across countries. We found all migratory groups showing higher Openness and Conscientiousness than non-migrants. Contrary to some previous accounts, Extraversion, Agreeableness and Neuroticism showed virtually no effects on migration distance. Our novel contribution is in the inclusion of a previously unstudied international type of migration (iv), a direct comparison of personality scores across all migration types, and a replication of results for the USA and Canada. This report highlights the mechanisms of how one's transitory psychological configuration can assist, inhibit or simply accompany the act of migration.

Keywords: immigrants, migration, migration distance, personality, Big Five

1.3 Introduction

Human migration is a phenomenon that has always affected and shaped socio-demographic dynamics. It manifests itself in diverse forms, e.g., as a permanent or temporary change of residence within the same local area, from a rural to an urban area or vice versa, or from one state or country to another. Reasons for migrations range from macro-level causes, i.e., differentials in work demand and supply between countries, to micro-level causes, such as individual values, desires and expectancies (Davin, 1999). While the former is mainly contingent on cultural, economic and political variables, the latter stem from the psychology of the migrating individual (Boneva & Frieze, 2001; Oishi, 2010).

Existing literature couched in the framework of migration psychology (Fawcett, 1985) argues that the prominence of particular personality traits predicts greater or smaller distances of migration, e.g., no migration, migration within a narrow local area, across state lines or across countries (Jokela et al., 2008; Jokela, 2009; Murray et al., 2005). For example, Jokela, Elovainio, Kivimäki, and Keltikangas-Järvinen (2008) considered a sample of 1,733 Finnish individuals and asked whether three temperament traits—sociability, emotionality, and activity—affected their migration propensity and migration distance. They found that high sociability was generally associated with within-country rural-to-urban migration but was also predictive of greater migration distances. Higher emotionality, on the other hand, predicted shorter migration distances. Additional longitudinal work confirmed the influence of personality traits on residential mobility and migration distance. In a sample of 3,760 individuals, Jokela (2009) examined the probability of migrating within or across the US states. Migration both within- and between-states came with higher openness and lower agreeableness scores; higher levels of extraversion, on the other hand, were only associated with within-state migration.

Migrant individuals are likely to face different challenges based on how far they move from their previous residence. Longer distances of migration typically come with different climatic, linguistic and demographic landscapes, i.e., factors that are shown to be related to personality traits (Allik & McCrae, 2013; McCrae, Terracciano, Realo, & Allik, 2007). Furthermore, the relationship between personality traits, the decision to migrate and migration distance is also explicable in terms of people's preferences towards environments in which they choose to reside. For example, the association between higher levels of agreeableness and low migration propensity suggests that more agreeable individuals tend to build stronger community ties (Lounsbury et al., 2013) and are therefore less inclined to change residence. This line of reasoning is supported by other studies that point out the association between low residential mobility and high social cohesion (Kan, 2007). The fact that agreeable people are more likely to stay in the same residential location than non-agreeable people may be in part explained by their higher affiliation motive, and therefore individuals who stay in the same neighborhood may be perceived as being more trustworthy and helpful than those who move. In sum, there is ample evidence that personality traits are predictive of whether and how individuals migrate.

The present report furthers this line of research in two ways. We examine a much larger sample than any preceding study and we add international migration to the better studied types of migration, i.e., within a state and across states in the same country. This broader scope of analysis is possible thanks to the personality and demographic data from a large-scale survey OutOfService.com (https://www.outofservice.com). This dataset includes Big 5 personality scores and a range of demographic data for millions of individuals, including the location of their current and previous residence as well as years after immigration (where applicable). Respondents were all volunteers who participated from 1997 to 2015. We analyzed differences

in levels of the Big 5 personality traits, aggregated separately for a range of migration distances (from non-migration to international migration). Thus, our study assesses which traits, if any, predict different migration distances.

1.4 Methods

The OutOfService dataset contains 9,328,610 respondents and their demographic data, which include their age, gender, race, as well as the country, state or province, and zip code of their current and/or previous residences. A change in residence address at any level was quantified to evaluate whether migration has taken place and whether it crossed the state or country lines: see categories of migration below. Personality traits for all respondents were assessed through the Big Five Inventory (John & Srivastava, 1999), which consists of 44 short statements designed to measure an individual on the Big Five Factors of personality dimensions. Respondents indicated the extent to which they agreed with each statement, on a scale from 1 (Disagree strongly) to 5 (Agree strongly).

We constrained the original sample to respondents with either the USA or Canada as the current country of residence, reducing the poll to 1,674,852 respondents (1,580,895 from the USA; 90,238 from Canada). We additionally excluded participants who left out information necessary for establishing their distance of migration, e.g., their previous and current zip code, state or country of residence. Using the demographic information provided by respondents, we were able to determine the distance of migration that some respondents went through. The dataset only identifies one previous place of residence: thus, we cannot capture cases when an individual undergoes multiple migrations. We identified four groups based on migration distance: participants who have (i) never changed their residence, (ii) moved within a state or a

province, (iii) moved across states or provinces, or (iv) moved to the USA or Canada from another country. Numbers of responders in each group are reported in Table 1, for Canada, the US and in total. Our use of these secondary data was approved by the McMaster Research Ethics Board (2018-089).

Groups (i-iv) constituted an independent ordinal variable with four categories representing populations currently residing in the US or Canada. Dependent variables were five personality traits from the Big Five Inventory, considered separately.

Table 1. Number of respondents by migration group

	Canada	USA	Total
Non-migrants	40,221	683,113	723,334
Moved within state or province	36,368	510,374	546,742
Moved across states or provinces	8,823	383,074	391,897
Moved across countries	4,784	4,292	9,076

1.5 Results

In this report we focused on the association between specific personality traits with the tendency to migrate both within and across national borders. We assessed what traits predicted different distances covered by migrants by aggregating levels of each personality trait at every migration category (i-iv). Then categories (ii) to (iv) were compared with all lower categories of this ordinal variable: for instance, we compared personality traits of international migrants (category iv) to migrants across state lines (iii), migrants within a state (ii) and non-migrants (i).

In these comparisons, each of the personality traits (i.e., openness to experience, conscientiousness, extraversion, agreeableness and neuroticism) was considered separately.

Given the massive size of our sample, standard p-values of a statistical test applied to the data (e.g., the two-sample t-test or Wilcoxon test) were not diagnostic. That is, virtually all p-values were much below the nominal 5% level of significance even after the family-wise correction for multiple comparison was applied. Therefore, we used Cohen's d (the standardized difference of means) to measure the effect size, i.e., the magnitude of a difference in a mean personality trait's score between each pair of samples.

The results were consistent for both the US and Canada (see Table 2): effect sizes that reached or exceeded the accepted level for a small effect size ($d \geq 0.2$) are shown in bold. All effects of migration type on personality traits ranged between negligible to small ($d < 0.27$). Yet, what the estimates of effect size indicated was a hierarchy of the predictive power of personality traits. Openness and conscientiousness were found to have the strongest effects on migration patterns. In fact, any move came with a greater level of openness and conscientiousness than found in individuals who did not migrate. Therefore, we treated the non-migratory group as the baseline to which compare the rest of the migratory groups. Specifically, the greatest differences in openness and conscientiousness were found between non-migrants and across- states migrants (openness: d=-.21 for USA, d=-.26 for Canada; conscientiousness: d=-.27 for USA, d=-.25 for Canada); the second group by largest effect size is that of responders who migrated across countries (openness: d=-.19 for USA, d=-.21 for Canada; conscientiousness: d=-.21 for USA, d=-.19 for Canada), followed by responders who migrated within states (openness: d=-.09 for USA, d=-.12 for Canada; conscientiousness: d=-.23 for USA, d=-.21 for Canada). No substantial differences in any other personality traits were observed between any pair of migratory groups in

the USA or Canada, regardless of whether they moved within a state, across states or between countries.

1.6 Discussion

We used a sample of 1.6 million American and Canadian migrant and non-migrant responders to determine whether certain personality traits can predict both people's tendency to migrate and the distance covered by the migration. We summarize our findings below and pit them against the current theoretical and empirical accounts.

This exploratory analysis confirmed and refined the hypothesis that different personality traits are associated with different types of migration, on a local, national and international scale. Specifically, we found that people who changed residence at some point of their life are more open and conscientious than people who did not move; this is in alignment with findings by Paulauskaitė et al. (2010) and it applies to migrants who moved within and between states and also across countries. For both Canada and the US, the greatest difference in openness and conscientiousness was shown between individuals who never moved and individuals who moved across states or provinces within the same country. Also, while migration at any distance elicited noticeable differences with the non-migratory group, the differences observed between those migratory groups were relatively small. And finally, while openness to experience is reliably associated with migration propensity both in this analysis and in the literature, our results for the other traits confirmed findings from some studies but not others. What follows is an overview on each personality trait, grouped by openness, conscientiousness and all the other traits, some of which did not match with previous findings in the literature.

Openness

The association between openness and migration propensity seems to stay consistent among the literature. It appears, not surprisingly so, that being open-minded is the main factor linked to the tendency to migrate. The fact that open-minded people are more likely to undertake a change of residence is not surprising. Adaptability is part and parcel of openness and personality trait, and individuals who perceive themselves as being more adaptable can show a greater inclination towards (or a lesser aversion to) adapting to physical, economic, and social challenges that migration might pose. Thus, fitting in with others' personality has shown to have effect on one's self esteem, especially for openness, agreeableness and conscientiousness (Bleidorn et al., 2016). We also found that greater openness does not necessarily come with a greater distance of migration. In fact, while moving out of a state requires more openness than moving within a state, moving across countries does not require more openness than moving within a country.

Conscientiousness

Similarly to openness, moving across countries does not require more conscientiousness than moving within a state or within a country. In fact, people who move out of a country score lower in conscientiousness than those who moved either within a state or across states. Thus, the effects of openness and conscientiousness appear to be categorical: higher levels of these traits lead to a higher propensity of *any* migration as compared to non-migration, regardless of its distance.

Discrepancies with existing literature

We did not replicate some of the existing findings from the literature. For example, we did not find any relation between agreeableness, extraversion or neuroticism and any type of migration considered in this analysis. Failing to detect any effect for agreeableness does not support the claims in Jokela (2009). Moreover, while other studies found extraversion to be associated both with the likelihood of national and international migration (Camperio Ciani et al., 2007; Jokela et al., 2008; Silventoinen et al., 2008), we found extraversion to not be a meaningful predictor. Finally, Jokela (2009) found no effect for neuroticism, while both Silventoinen et al. (2008) and Jokela et al. (2008) showed a positive relation between this trait and the tendency of changing location of residence. We observed that higher levels of neuroticism were linked with international migration, but the effect was negligibly small. Given our sample size, it is highly likely that the null effects we observed are representative of the US and Canadian populations at large.

Summary

We report new empirical evidence from a very large-scale dataset regarding the relationship between personality traits, propensity for migration and distance of migration. As mentioned above, effect sizes of comparisons between these four populations are relatively small, with values that range from 0.001 to 0.27. This suggest the presence of a significant overlap between our groups, e.g., most of the people who moved across countries do not differ significantly from those who never moved in their personality traits. Despite the small effect sizes, our analysis also provides a few novel contributions to the migration psychology literature. First, this is the first study on personality and distance of migration that compared between each other personality traits associated with all migration types. Second, we included a novel type of international migration. Third, our results challenge some of the previous findings in the

literature. Given the centrality of migration in every place and time of human progress, this brief

report puts on display the process of how one's transitory psychological configuration can assist,

inhibit or simply accompany the act of migration.

Table 2. Measures of effect size for each pair of groups. Marked in bold are Cohen's d values which exceed the threshold for the small effect (|d| > 0.2).

USA	Trait	Migration within states	Migration across states	Migration across countries
	ope	-0.09	**-0.21**	-0.19
	con	**-0.23**	**-0.27**	**-0.21**
Non-migration	ext	0.02	-0.03	-0.014
	agr	-0.1	-0.06	-0.01
	neu	0.06	0.12	0.148
	ope		-0.12	-0.1
	con		-0.03	0.02
Migration within states	ext		-0.01	0.006
	agr		0.04	0.09
	neu		0.06	0.08
	ope			0.02
	con			0.06
Migration across states	ext			0.01
	agr			0.049

Canada		Migration within states	Migration across states	Migration across countries
	neu			0.025
Non-migration	ope	-0.12	**-0.26**	**-0.21**
	con	-0.19	**-0.25**	-0.19
	ext	0.01	0.02	0.003
	agr	-0.05	-0.02	-0.05
	neu	0.003	0.01	0.06
Migration within states	ope		-0.13	-0.08
	con		-0.06	0.004
	ext		0.01	-0.01
	agr		0.03	-0.001
	neu		0.01	0.05
Migration across states	ope			0.04
	con			0.06
	ext			-0.02
	agr			-0.03
	neu			0.04

1.7 References

Allik, J., Realo, A., & McCrae, R. R. (2013). Universality of the five-factor model of personality.

Bleidorn, W., Schönbrodt, F., Gebauer, J. E., Rentfrow, P. J., Potter, J., & Gosling, S. D. (2016). To live among like-minded others: Exploring the links between person-city personality fit and self-esteem. *Psychological science*, *27*(3), 419-427.

Boneva, B. S., & Frieze, I. H. (2001). Toward a concept of migrant personality. *Journal of Social Issues*, 57, 477–491.

Camperio Ciani, A. S., Capiluppi, C., Veronese, A., & Sartori, G. (2007). The adaptive value of personality differences revealed by small island population dynamics. *European Journal of Personality: Published for the European Association of Personality Psychology*, *21*(1), 3-22.

Davin D. (1999) Why People Migrate. In: Internal Migration in Contemporary China. Palgrave Macmillan, London

Fawcett, J. T. (1985). Migration psychology: New behavioral models. *Population and Environment*, *8*(1-2), 5-14.

John, O. P., & Srivastava, S. (1999). The Big Five trait taxonomy: History, measurement, and theoretical perspectives. *Handbook of personality: Theory and research*, *2*(1999), 102-138.

Jokela, M. (2009). Personality predicts migration within and between US states. *Journal of Research in Personality*, *43*(1), 79-83.

Jokela, M., Elovainio, M., Kivimäki, M., & Keltikangas-Järvinen, L. (2008). Temperament and migration patterns in Finland. *Psychological Science*, *19*(9), 831-837.

Kan, K. (2007). Residential mobility and social capital. *Journal of Urban Economics*, 61, 436–457.

Lounsbury, J. W., Loveland, J. M., & Gibson, L. W. (2003). An investigation of psychological sense of community in relation to big five personality traits. *Journal of Community Psychology, 31*, 531–541.

McCrae, R. R., Terracciano, A., Realo, A., & Allik, J. (2007). Climatic warmth and national wealth: some culture-level determinants of national character stereotypes. *European Journal of Personality: Published for the European Association of Personality Psychology*, *21*(8), 953-976.

McCrae, R. R., Terracciano, A., Realo, A., & Allik, J. (2007). On the validity of culture-level personality and stereotype scores. *European Journal of Personality: Published for the European Association of Personality Psychology*, *21*(8), 987-991.

Murray, G., Judd, F., Jackson, H., Fraser, C., Komiti, A., Hodgins, G., ... & Robins, G. (2005). The five factor model and accessibility/remoteness: Novel evidence for person–environment interaction. *Personality and Individual Differences*, *39*(4), 715-725.

Oishi, S. (2010). The psychology of residential mobility: Implications for the self, social relationships, and well-being. *Perspectives on Psychological Science*, *5*(1), 5-21.

Paulauskaitė, E., Šeibokaitė, L., & Endriulaitienė, A. (2010). Big five personality traits linked with migratory intentions in Lithuanian student sample. *International journal of psychology: a biopsychosocial approach, 2010,[Vol.] 7, p. 41-58*.

Silventoinen, K., Hammar, N., Hedlund, E., Koskenvuo, M., Rönnemaa, T., & Kaprio, J. (2007). Selective international migration by social position, health behaviour and personality. *European Journal of Public Health*, *18*(2), 150-155

2.0 Influence of Immigration on Personality Traits

Davide Gentile[1], Constance Imbault[1], Samuel Gosling[2],

Jason Rentfrow[3], Jeff Potter[4] & Victor Kuperman[1]

[1]McMaster University

[2]University of Texas at Austin

[3]University of Cambridge

[4]Enveritas

2.1 Preface

This article has been prepared to be submitted as a brief report to the Journal of Research in Personality. The data were collected by Dr. Samuel Gosling, Dr. Jason Rentfrow and Dr. Jeff Potter, who agreed to share it with us. The conceptualization of this report was conducted by myself and Connie Imbault, in collaboration with Dr. Kuperman. I explored, cleaned and analyzed the data under the supervision of Dr. Kuperman. The statistical analysis was completed by myself, Constance Imbault and Dr. Kuperman. The article was written by myself, with edits from Dr. Victor Kuperman and feedback from Constance Imbault and Dr. Jason Rentfrow.

2.2 Abstract

Openness is the only Big Five personality trait consistently associated with the propensity of migration. We used a large-scale data set with personality scores and demographic information (N = 1,199,546) to examine how the immigration status of an individual affects their levels of openness, and how this relationship is modulated by other known determinants of personality traits: age and gender. Our study provides new evidence that age and gender interact with openness of immigrants and non-immigrants in divergent ways. Using Generalized Additive Models, we found both male and female immigrants to be less open than non-immigrants in early adulthood (21-27 y.o.). We also found that male immigrants show higher levels of openness than non-migrants in the age range of 35-55 y.o. Conversely, adult female immigrants show no difference from non-immigrants in levels of openness. The novelty of our study is in analyzing a very large sample, the inclusion of social variables such as age and gender, and their interaction with immigration status as predictors of openness. These findings help shed light on the psychological effects of immigration and consider how these effects are modulated by people of different ages and gender.

Keywords: immigrants, migration, personality, Big Five, openness, age, gender

2.3 Introduction

Immigrants and immigration are subjects of study in many disciplines. Within this interdisciplinary field of investigation, the motivations, values and aspirations of individuals who decide to migrate indicate the centrality of psychology and the social sciences in uncovering these questions. The aim of this report is to provide new evidence on the role of migration in shaping an individual's personality. In the following introductory sections, we first illustrate the relationship between immigration and personality traits, then report what is generally known about personality development and finally explain why, out of the Big Five traits, we decided to focus specifically on openness to new experience.

Immigration and personality

Researchers in the field - labeled either 'migration psychology' (Fawcett, 1985) or 'psychology of immigration' (Berry, 2001) - study psychology and behavior of the migratory individual. One commonly asked question in this field is whether the intention to migrate or the act of migration bring with them a change in a personality of an individual. A standard instrument used in such studies is a measurement of the Big Five personality traits: openness, conscientiousness, extraversion, agreeableness and neuroticism (Barrick & Mount, 1991; Goldberg, 1990). Differences in distributions of personality scores in each of the five dimensions have been shown to estimate one's propensity to migrate (Camperio Ciani et al., 2007; Silventoinen et al., 2008), the distance covered by the migration (Jokela 2008; 2009) and the psychological determinants of migration intentions (Canache, Hayes, Mondak & Wals, 2013; Fouarge, Ozer & Seegers, 2016; Paulauskaitė et al., 2010).

Personality development

      Although the personality traits framework is commonly adopted, there is no solid

agreement on what personality traits are and how they develop. The literature acknowledges two

main perspectives that are thought to explain the nature of personality traits. Following a

perspective that is exclusively biological, one's psychological asset is "insulated from the direct

effects of the environment" (McCrae & Costa, 1999, p. 144). According to this view, personality

traits are malleable during childhood and they reach maturity by the age of 30; because of the

steady asset that personality is thought to become, this theory is sometimes referred to as *the*

plaster theory (Srivastava et al., 2003). On this view, a certain configuration of personality traits

may grant an adult individual with a higher propensity to migrate, but the act of migration is not

expected to affect their personality. The opposing view, generally called a *contextual* or

environmental perspective, argues that personality is subject to changes even over the course of

adulthood (e.g., Haan et al., 1986; Helson, Mitchell, & Moane, 1984; Neugarten, 1972) as a

function of adaptation to changing environment. This perspective opens a discussion about what

factors influence or modulate personality changes over time (potentially, including migration),

and is gradually gaining support. For example, many replicated findings indicate the presence of

age differences in the Big Five trait domain, with the most sensible period being from late

childhood through middle age (Roberts & Mroczek, 2008; Soto, John, Gosling & Potter, 2011).

In fact, this phase generally shows positive trends in conscientiousness and agreeableness, and a

gradual decline in neuroticism (Allemand, Zimprich & Hendriks, 2008; Donnellan & Lucas,

2008; Terracciano, McCrae, Brant, & Costa, 2005). Differences in personality traits are also

found between genders (Helson, Pals, & Solomon, 1997; Weisberg, DeYoung, & Hirsh, 2011;

Wink & Helson, 1993). Gender differences in the Big Five traits are generally intended as which gender - on average - scores higher on a given trait. For example, women are usually found to be more agreeable than men, and to a less extent, more neurotic (Feingold, 1994; Costa et al., 2001).

We propose that migration might be one of the factors modulating people's changes in personality over a lifetime. In fact, while setting in a new place, immigrants are likely to be affected by numerous disadvantages, from underemployment to poor social support and lack of access to services (Simich, Beiser, Stewart, & Mwakarimba, 2005). This sudden disturbance could cause changes in their patterns of thoughts and behaviour: the original use of the expression 'cultural shock' was indeed referred to the sense of uncertainty and aversion in an unfamiliar cultural environment (Zhou, Jindal-Snape, Topping & Todman, 2008). As a matter of fact, the borders that migratory individuals try to cross are not only geographical, but also linguistic and cultural (Dewaele & Stavans, 2014). In this scenario, the psychological tools that one has available can help determin the outcome of their cultural assimilation. In other words, immigrants' capacity to adapt to a new environment becomes a meaningful predictor of successful integration. Additionally, some of the challenges faced by immigrants seem to address aspects that are specifically related to openness to experience. In fact, by finding themselves beyond their comfort zone, immigrants are almost required to enhance aspects of their personality that are characteristic of people who generally score high in this Big Five dimension. For example, students scoring high in openness tend to focus "on meaning, possibilities, and usually accept constant change" (Ehrman 2008, p. 66), all concepts that need to be embraced when setting in a new place. Another example involves the social isolation that immigrants are likely to face, especially in the first phases of their migration (Hagan, 1998; Grieco, 1998). Rebuilding disrupted social networks requires motivation and proactive participation in the

community; for immigrants, this means giving up the sense of comfort originated from the (more predictable) environment that they are accustomed to – and act as they are open individuals, even if originally they were not. In the next section, we focus on openness to experience and how it is shown to engage with migration propensity and some other aspects of integration.

Openness

A variety of factors, spanning from the economic to the sociological ones, can influence one's propensity to migrate. As recent studies indicate, personality differences may also play a role in such decision; in fact, openness to experience seems to be a particularly good predictor of migration propensity (Camperio Ciani et al., 2007; Canache et al., 2013; Gentile et al., 2019; Jokela 2008, 2009; Silventoinen et al., 2008), suggesting that open-minded people, willing to face unfamiliar challenges, are more likely to migrate to different places of residency. Afterall, individuals high in openness proactively seek new experiences and appreciate the unfamiliar, while those who score low on this dimension are more conventional and prefer dealing what they already know. Openness has also been found to be a significant predictor of foreign language learning achievement (Ozanska & Dewaele, 2012, Verhoeven & Vermeer 2002). Ozanska and Dewaele show that participants who score higher in openness and self-esteem are more likely to seek authentic interactions in their second language. This suggests that immigrants' openness is a meaningful source of individual difference when it comes to successful second language development and consequently the number and nature of their interactions.

This study

In this study we treat openness scores as a function of immigration status and investigate the modulating role of other well-known determinants of personality traits: age and gender. As mentioned above, the relation between personality and age is non-linear, i.e., the rate of increase or decrease of a given trait's score changes as age changes. For this reason, we believe the trend we analyze in our study to be likely to be better modeled by a nonlinear function.

The data used in our analysis is originated from the online source OutOfService.com (https://www.outofservice.com), where millions of respondents have been providing personality and demographic data voluntarily since 1997. In this large-scale dataset (original N=9,328,610), personality scores were assessed based on the Big Five Inventory (John & Srivastava, 1999): respondents indicated the extent to which they agreed with each of the 55 statements, on a scale from 1 (Disagree strongly) to 5 (Agree strongly); this provided averaged individual scores for each of the Big Five dimensions. This cross-sectional dataset gives us access to personality scores of an individual at a given age and a given immigration status (immigrant or not). Thus, we do not have direct evidence of individual personality before migration or one's propensity to migrate. Thus, we consider point estimates of immigrants' personality as a juxtaposition of their personality configuration before migration and the contextual influence of migration on personality traits. Our question is whether immigration have an effect on openness levels and if this influence is modulated by two major contextual variables. We illustrate the methods below.

2.4 Methods

The dataset collected at <u>outofservice.com</u> reports personality scores of respondents and their demographic data, which include their age, gender, race, as well as the country, state/province, and zip code of their residences in the present and in the past. We examined change in residence address to evaluate whether migration has taken place at the country level. We excluded participants who left out information necessary for establishing whether a migration occurred within or across the national boundaries. We also constrained the original sample to respondents with either the USA or Canada as the current country of residence. Using the demographic information provided by respondents we were able to divide Americans and Canadians who never changed country of residence from the people who migrated into one of these two country at some point in their life. We decided to only include Canada and US in our analysis because the majority of respondents resulted as coming from one of these two countries. The exclusion criteria reduced the poll to 1,199,546 respondents: 1,138,416 non migrants (388,268 females) and 61,130 migrants (23,481 females). We adopted a pseudo-longitudinal approach to monitor temporal changes in personality: respondents provided their age only once at the time when the survey was taken. Numbers of responders used in our analysis are reported in Table 1, classified by immigration status, country and gender. Our use of these secondary data was approved by the McMaster Research Ethics Board (2018-089).

Table 3. Number of respondents by immigration status, country and sex.

	Canada	USA	Total
Immigrants	12,716	48,414	61,130
	(5,038 females)	(18,443 females)	(23,481 females)
non-migrants	61,109	1,077,307	1,138,416
	(21,033 females)	(367,235 females)	(388,268 females)

To investigate the effect of immigration status to the openness scores of respondents, we used generalized additive mixed-effects model (GAMM). As mentioned in the introduction, we expect our approach to result in non-linear functional forms, which GAMM is particularly suitable to model. Models were fitted in the programming environment R (version 3.5.0), using the package *mgcv*, version 1.8-23 (Wood, 2006). Given the size of our sample, we used the function *bam*, which is designed for datasets exceeding tens of thousands of data points (Wood, Goude and Shaw, 2015).

The summary information for each of the models is reported in Table 2, 3 and 4, broken down by parametric coefficients and smooth effects. The parametric coefficients report the simple main effect of age, while the non-linear effects and interactions (age by immigration status, age by immigration status for males and age by immigration status for females) are given in smooth terms. To illustrate the estimated effect, we used the R package *itsadug* (version 2.3), whose function *sim.ci* allowed us to simultaneously construct the confidence interval (CI): in doing so, we were able to identify the key ages where significant differences in openness occur, i.e., both early and late adulthood.

2.5 Results

This section reports the results of our analysis and interprets empirically the influence of immigration on openness development of respondents. The questions were whether immigrants and non-immigrants differed in their trend of openness as a function of age, and if this difference was reflected when controlling for gender. In other words, we were interested in knowing

whether aging or gender affected immigrants' openness scores differently than non-immigrants' openness scores.

We first assessed how immigrants' and non-immigrants' levels of openness develop differently with age and then focused on the separate trends showed by male and female respondents. The results of our two analyses are discussed separately below.

2.5.1 Aging

For modeling potentially different trends over time for immigrants and non-immigrants, we set up a GAMM model with immigration status as the grouping predictor. Our response variable (openness score, ranging 0-5) was related to the predictor of interest *age*, for both immigrants and non-immigrants. By including the intercept, the model is more flexible in capturing any intercept differences between the two groups of interest. Table 2 reports the summary information of the model.

Both smooth terms (non-linear effects of age for immigrants and non-immigrants) in Table 2 are highly significant (p < 0.001) suggesting that age has an effect on openness in both populations. This however does not allow us to establish whether the two curves are different from each other. We therefore calculated the estimated difference in age effect between the two conditions, i.e., immigration and non-immigration. In this way, we were able to isolate the age windows where significant differences in openness take place and the sign of the difference (positive or negative). We found indeed that aging affects immigrants and non-immigrants differently, see Figure 1. As Figure 1 shows, in both populations the overall trend is for individuals to report higher scores of openness over the life span, and especially in mature adulthood (over 35 y.o.). A

comparison between the curves additionally showed that at younger ages immigrants tend to score lower in openness than their non-immigrant peers. The window in which the contrast was estimated as statistically significant was 21-27 years of age. This window is followed by a plateau (27-35 y.o.), in which the openness levels of the two populations do not differ in a statistically reliable way. However, the pattern of results appears to change radically at later stages in life, as immigrants show higher openness scores than non-immigrants in adult years. The age range where immigrants score reliably higher in openness than non-immigrants was estimated as 35-53 years of age.
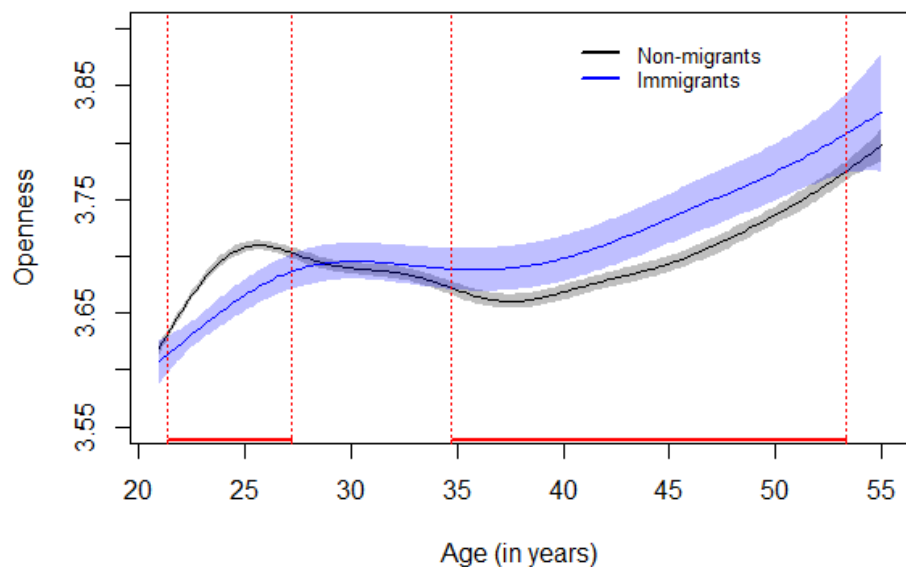

Figure 1. Different personality trends for immigrants and non-immigrants. While at younger ages (21-27 y.o.) immigrants score lower than non-immigrants in openness, during adulthood (35-53 y.o.) immigrants become more open than non-immigrants.

2.5.2 Gender

For modeling the different age-related trends for male and female immigrant and non-immigrant respondents over time, we set up a GAMM model with immigration status as the grouping predictor - similarly to the analysis above - but we additionally studied patterns in our data broken down by gender. We fitted a GAMM model to each subset of the data: Our response variable was *openness*, operationalized as openness scores in the Big 5 test, while age and immigration status were predictors of interest. Table 3 and Table 4 report the summary information of the models for male and female respondents, respectively.

One of our main findings was that male and female immigrants differ in their trends in personality, especially as they get older. The male population (Figure 2) closely followed the same cross-over tendency that we reported for the population at large. At least partly, this is because male respondents were better represented in both the non-immigrant (65%) and the immigrant (62%) samples that we consider. Specifically, non-immigrants showed a steady increase in openness across the entire age range under consideration, while immigrants were reliably less open than non-immigrants in the age range of 21-29 y.o. and reliably more open between ages 35 and 55.

The analysis of the female subset showed a very different pattern, see Figure 3. Unlike the cross-over non-linear interaction observed in Figure 2, the curves representing model-estimated openness of immigrant and non-immigrant women across the lifespan are very similar to each other in shape but shifted over the age axis. Both immigrant and non-immigrant women experience an initial increase in openness followed by a decrease (an inverse-U shape) in their 20s, which is then followed by a sustained increase in openness until the end of the age range under consideration. The difference between the two curves is that the one representing non-

immigrant women peaks earlier (around 25 y.o.), while that for immigrant women peaks at around 28 y.o. This difference is statistically reliable between ages 22-26 and is generated by a shift in otherwise overlapping curves.
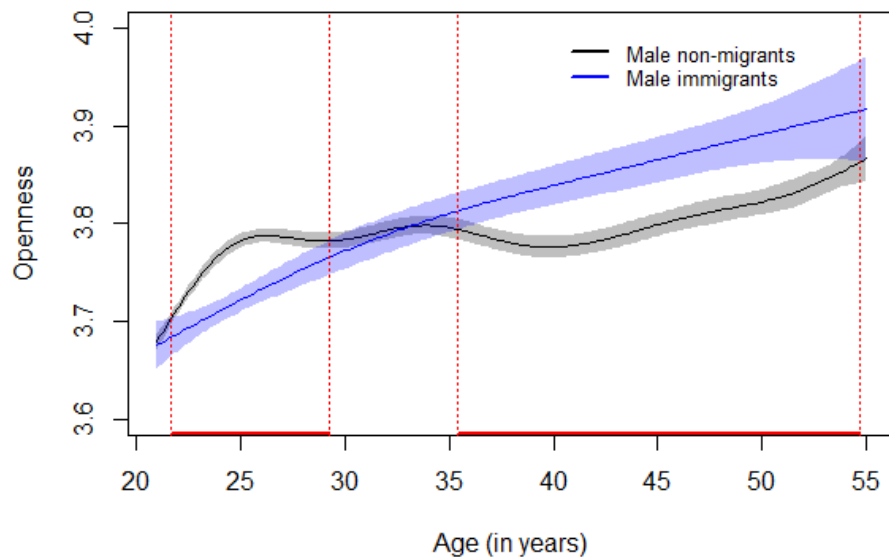

Figure 2. Young male immigrants (21-29 y.o.) scored lower than their non-migrant peers in openness. This tendency changes at older ages, where adult male immigrants (35-55 y.o.) show a consistent increase in openness scores and overreach their adult male non-migrant peers.
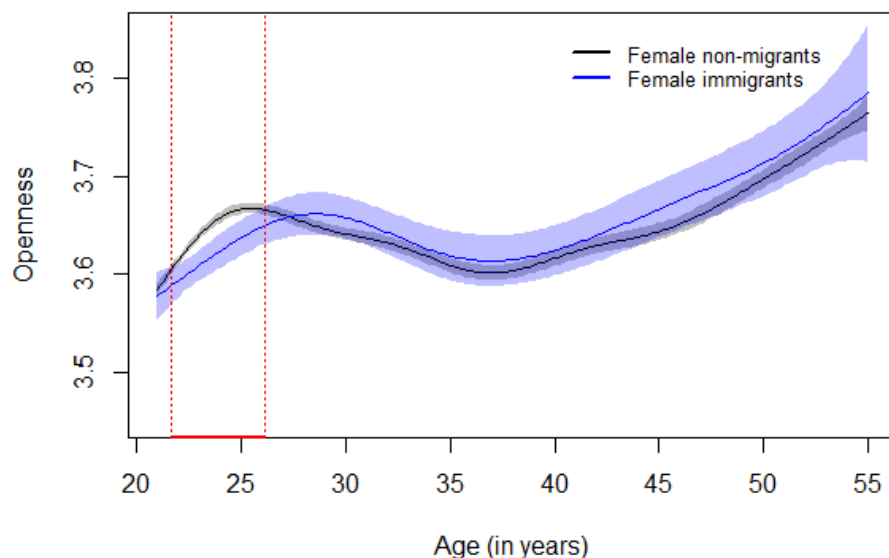
Figure 3. Female immigrants are less open than female non-immigrants in early adulthood (21-26 y.o.), but there are no significant differences during late adulthood.

2.6 Discussion

The goal of this study is to provide new evidence on the role of migration in modulating an individual's personality. Our exploratory analysis contributes to the literature on migration psychology by examining the effect of immigration status on one Big Five trait that is consistently associated with migration propensity: openness to experience. We used a sample of 1,199,546 American and Canadian immigrant and non-immigrant responders to determine whether immigrants experience different trends of personality development than non-immigrants. In our investigation we additionally divided respondents by age and gender, which as stated in the introduction, are considered variables with a major role in determining outcomes of personality traits. We summarise and interpret our results below.

We found that aging has an effect on both immigrants' and non-immigrants' openness, and that the magnitude of this effect changes at different age windows. Generally, our results show that during early adulthood all immigrants – female and male - are less open than their non-immigrant counterparts. Specifically, this contrast was statistically significant between the age range of 21-27 years. One possible explanation is related to the fact that immigrant responders in their early 20s could be students or foreign workers whose legal staying in Canada or US is limited to the duration of their contracts. In fact, although they might share some characteristics with their non-immigrant peers, such as youth and student/worker status, they are allowed in Canada and US only temporarily, and are not considered permanent residents (Government of Canada, 2019). Berry, Kim, Minde, & Mok (1987) link the transitionality of Canadian immigrants' legal sojourn with their unwillingness to establish supportive relationships among themselves or in their host community. We suggest that the lack of supportive networks could be one factor related to the lower levels of openness reported by young male and female immigrants in Canada and the US.

However, after a period of no difference in openness between the two groups (age range of 27-35 years), our results also show that immigrants become more open than non-immigrants during the age range of 35-53 years. During the non-significant window between 27-35 years of age, people generally become more conscientious, more agreeable and less neurotic (Allemand, Zimprich & Hendriks, 2008; Donnellan & Lucas, 2008; Roberts & Mroczek, 2008; Soto, John, Gosling & Potter, 2011; Terracciano, McCrae, Brant, & Costa, 2005). One could argue that during this stage of life, both immigrants and non-immigrants find themselves in the process of negotiating important life achievements, such as pursuing a career or establishing authentic relationships (Hogan & Roberts, 2004). Moreover, although further adjustments are possible,

one's levels of openness are likely to not undertake drastic changes across early adulthood and middle age (Soto, John, Gosling & Potter, 2011). Following this perspective, we propose that this is part of the reason why immigrants and non-immigrants show no significant differences in openness during the age window resulted from our data.

During the following age window (35-53 years), immigrants and non-immigrants differ in their openness levels, with immigrants scoring significantly higher. It might be the case that immigrants who went through some problematic implications of immigration (refer to the Introduction) have learnt to navigate unfamiliar waters and change in response to environmental demands. These skills may lead immigrants to develop higher curiosity and imagination and to be more unprejudiced than their non-immigrants counterparts – all features associated with openness to experience (DeYoung, 2014; DeYoung et al., 2014).

We also found that men and women vary in their openness as a function of immigration status, and especially in mature adulthood, after 30 y.o. Specifically, we found young male immigrants, whose openness scores are generally lower than those of their peers, to experience a progressive increase in openness levels with age: this increase lead male immigrants to not only reach the same openness scores of male non-immigrants, but also to exceed them starting approximately from the age of 35. This is in line with what we found in the sample that was not divided by gender, therefore the same explanation we proposed for age differences for immigrants are also applicable to male immigrants specifically. However, the trend is different for adult female immigrants. In fact, it appears that starting from young adulthood, in particular from the age of 26 years, female immigrants do not differ significantly from female non-immigrants in openness trajectories. The pattern suggests that the difference in openness between older and younger immigrant women is smaller than that between older and younger immigrant

men. In addition, while male immigrants and non-immigrants never really align in their openness scores, age moderated the difference in openness in immigrant and non-immigrant females such that the difference between the two groups was smaller at higher ages. In fact, it seems that female immigrants and non-immigrants become similar in openness as they get older. This pattern most likely indicates that immigration status modulated openness in older immigrants differently, based on their gender. It might be worth furthering this line of inquiry by examining gender differences in openness at the facet level. Since differences in openness are mostly negligible in adulthood across genders (Soto, John, Gosling & Potter, 2011), one possible interpretation of this finding might be that older women go through similar psychological processes regardless the immigration status. In fact, Orth et al. (2012) suggest that women improve the perception of their own intelligence as they get older, probably reflecting increases in self-esteem or self-confidence. Thus, since intelligence and self-esteem are both facets of openness in the Big Five domain, immigrant and non-immigrant women might align in their openness due to the change of their own self-perception. Given the gender differences found in openness at the facets level (Weisberg, DeYoung, & Hirsh, 2001), this could also explain why we do not observe the same trend for older male respondents.

This study also presents a few limitations. First, the information available in our data was given by respondents only once, at a given age and a given immigration status. Therefore ours was a pseudo-longitudinal approach: we did not observe individual personality shifts from pre-departure levels to post-departure levels. More longitudinal analyses could shed light on individual differences in personality development among immigrants. Moreover, we were not able to detect whether immigrants have moved more than once before answering the questionnaire. Thus, we treated respondents who moved only once the same as respondents who

moved multiple times. This generates further questions, such as: is the effect of migration on openness stronger after the first migration than after the following migrations? How does the effect of age modulate openness scores when respondents faced multiple changes of residence?

Second, the Outofservice questionnaire is only available in English. Personality differences have been found across cultures (Costa Jr, Terracciano, & McCrae, 2001; Eaves, Eysenck, & Martin, 1989), and extending our research questions to a more multicultural sample would allow to establish whether the personality trends that we reported for Canadian and American immigrants replicate in different countries.

In conclusion, this analysis aims at assessing the nature of a complex interaction among immigration, personality and two major contextual variables. In sum, our findings suggest that migration may influence personality changes of people of different age and gender. We report that male and female immigrants tend to be less open than non-immigrants during early adulthood (21-27 y.o.), that male immigrants reach higher levels of openness than non-immigrants after middle age (35-55 y.o.), and that adult female immigrants show no difference from their non-immigrants peers in levels of openness. With the present-day availability of large-scale datasets that are comprehensive of both demographic and psychological information, empirical findings like ours help shed light on the psychological effects of immigration and on how these effects are modulated by people at different ages and gender.

2.7 References

Allemand, M., Zimprich, D., & Hendriks, A. A. (2008). Age differences in five personality domains across the life span. *Developmental Psychology*, *44*(3), 758.

Barrick, M. R., & Mount, M. K. (1991). The big five personality dimensions and job performance: a meta-analysis. *Personnel psychology*, *44*(1), 1-26.

Berry, J. W., Kim, U., Minde, T., & Mok, D. (1987). Comparative studies of acculturative stress. *International migration review*, *21*(3), 491-511.

Berry, J. W. (2001). A psychology of immigration. *Journal of social issues*, *57*(3), 615-631.

Bodvarsson, Ö. B., & Van den Berg, H. (2013). *The economics of immigration*. New York, NY: Springer New York.

Borjas, G. J. (2014). *Immigration economics*. Harvard University Press.

Brettell, C. B. (2013). Anthropology of migration. *The Encyclopedia of Global Human Migration*.

Brettell, C. B., & Hollifield, J. F. (2013). Theorizing migration in anthropology: The social construction of networks, identities, communities, and globalscapes. In *Migration theory* (pp. 121-168). Routledge.

Camperio Ciani, A. S., Capiluppi, C., Veronese, A., & Sartori, G. (2007). The adaptive value of personality differences revealed by small island population dynamics. *European Journal of Personality: Published for the European Association of Personality Psychology*, *21*(1), 3-22.

Canache, D., Hayes, M., Mondak, J. J., & Wals, S. C. (2013). Openness, extraversion and the intention to emigrate. *Journal of Research in Personality*, *47*(4), 351-355.

Costa Jr, P. T., Terracciano, A., & McCrae, R. R. (2001). Gender differences in personality traits across cultures: robust and surprising findings. *Journal of personality and social psychology*, *81*(2), 322.

Dewaele, J. M., & Stavans, A. (2014). The effect of immigration, acculturation and multicompetence on personality profiles of Israeli multilinguals. *International Journal of Bilingualism*, *18*(3), 203-221.

DeYoung, C. G. (2014). Openness/Intellect: A dimension of personality reflecting cognitive exploration. *APA handbook of personality and social psychology: Personality processes and individual differences*, *4*, 369-399.

DeYoung, C. G., Quilty, L. C., Peterson, J. B., & Gray, J. R. (2014). Openness to experience, intellect, and cognitive ability. *Journal of personality assessment*, *96*(1), 46-52.

Donnellan, M. B., & Lucas, R. E. (2008). Age differences in the big five across the life span: evidence from two national samples. *Psychology and aging*, *23*(3), 558.

Eaves, L. J., Eysenck, H. J., & Martin, N. G. (1989). *Genes, culture and personality: an empirical approach*. Academic Press.

Ehrman, M. (2008). Personality and good language learners. *Lessons from good language learners*, 61-72.

Fawcett, J. T. (1985). Migration psychology: New behavioral models. *Population and Environment*, *8*(1-2), 5-14.

Fouarge, D., Ozer, M., & Seegers, P. (2016). *Personality traits and migration intention: Who bears the cost of migration*. Maastricht University, mimeo.

Freeman, G. P. (1995). Modes of immigration politics in liberal democratic states. *International migration review*, *29*(4), 881-902.

Gentile, D., Imbault, C., Gosling, S., Rentfrow, J., Potter, J., & Kuperman, V. (2019). (In Publication). Big Five Traits as Predictors of Migration Distance in Canada and the US**.**

Goldberg, L. R. (1990). An alternative" description of personality": the big-five factor structure. *Journal of personality and social psychology*, *59*(6), 1216.

Government of Canada. (2019). *Understand permanent resident status*. Retrieved from https://www.canada.ca/en/immigration-refugees-citizenship/services/new-immigrants/pr-card/understand-pr-status.html

Grieco, E. (1998). The Effects of Migration on the Establishment of Networks: Caste Disintegration and Reformation among the Indians of Fiji. Int Migr Rev 1998; 32:704–736.

Haan, N., Millsap, R., & Hartka, E. (1986). As time goes by: Change and stability in personality over fifty years. *Psychology and aging*, *1*(3), 220.

Hagan, J (1998). Social Networks, gender, and immigrant incorporation: Resources and constraints. Am Sociol Rev 1998; 63(1): 55–67.

Helson, R., Mitchell, V., & Moane, G. (1984). Personality and patterns of adherence and nonadherence to the social clock. *Journal of personality and social psychology*, *46*(5), 1079.

Helson, R., Pals, J., & Solomon, M. (1997). Is there adult development distinctive to women? In *Handbook of personality psychology* (pp. 291-314). Academic Press.

Hogan, R., & Roberts, B. W. (2004). A socioanalytic model of maturity. *Journal of Career Assessment*, *12*(2), 207-217.

Hollifield, J. F. (2000). The politics of international migration. *Migration theory: Talking across disciplines*, 137-85.

McCrae, R. R., & Costa, P. T., Jr. (1999). A five-factor theory of personality. In L. A. Pervin & O. P. John (Eds.), Handbook of personality: Theory and research (2nd ed., pp. 139–153). New York: Guilford Press.

Neugarten, B. (1972). Personality and the aging process. *The Gerontologist*.

John, O. P., & Srivastava, S. (1999). The Big Five trait taxonomy: History, measurement, and theoretical perspectives. *Handbook of personality: Theory and research*, *2*(1999), 102-138.

Jokela, M. (2009). Personality predicts migration within and between US states. *Journal of Research in Personality*, *43*(1), 79-83.

Jokela, M., Elovainio, M., Kivimäki, M., & Keltikangas-Järvinen, L. (2008). Temperament and migration patterns in Finland. *Psychological Science*, *19*(9), 831-837.

Orth, U., & Robins, R. W. (2014). The development of self-esteem. *Current Directions in Psychological Science*, *23*(5), 381-387.

Ożańska-Ponikwia, K., & Dewaele, J. M. (2012). Personality and L2 use: The advantage of

    being openminded and self-confident in an immigration context. *Eurosla*

    Yearbook, *12*(1), 112-134.

Roberts, B. W., & Mroczek, D. (2008). Personality trait change in adulthood. *Current directions*

    in psychological science, *17*(1), 31-35.

Silventoinen, K., Hammar, N., Hedlund, E., Koskenvuo, M., Rönnemaa, T., & Kaprio, J. (2007).

    Selective international migration by social position, health behaviour and

    personality. *European Journal of Public Health*, *18*(2), 150-155.

Simich, L., Beiser, M., Stewart, M., & Mwakarimba, E. (2005). Providing social support for

    immigrants and refugees in Canada: Challenges and directions. *Journal of Immigrant and*

    Minority Health, *7*(4), 259-268.

Soto, C. J., John, O. P., Gosling, S. D., & Potter, J. (2011). Age differences in personality traits

    from 10 to 65: Big Five domains and facets in a large cross-sectional sample. *Journal of*

    personality and social psychology, *100*(2), 330.

Srivastava, S., John, O. P., Gosling, S. D., & Potter, J. (2003). Development of personality in

    early and middle adulthood: Set like plaster or persistent change? *Journal of personality*

    and social psychology, *84*(5), 1041.

Terracciano, A., McCrae, R. R., Brant, L. J., & Costa Jr, P. T. (2005). Hierarchical linear

    modeling analyses of the NEO-PI-R scales in the Baltimore Longitudinal Study of

    Aging. *Psychology and aging*, *20*(3), 493.

Verhoeven, L., & Vermeer, A. (2002). Communicative competence and personality dimensions in first and second language learners. *Applied Psycholinguistics*, *23*(3), 361-374.

Weisberg, Y. J., DeYoung, C. G., & Hirsh, J. B. (2011). Gender differences in personality across the ten aspects of the Big Five. *Frontiers in psychology*, *2*, 178.

Wink, P., & Helson, R. (1993). Personality change in women and their partners. *Journal of Personality and Social Psychology*, *65*(3), 597.

Zhou, Y., Jindal-Snape, D., Topping, K., & Todman, J. (2008). Theoretical models of culture shock and adaptation in international students in higher education. *Studies in higher education*, *33*(1), 63-75.

3.0 Immigration and Personality in the Language of Facebook

3.1 Abstract

Immigrants and non-immigrants show differences throughout some dimensions of the Five Factor model of personality. Using demographic, psychological and textual data, we conducted two analyses to determine whether differences in personality of immigrants and non-immigrants are reflected in their language use on Facebook. In Study 1, we calculate what words are more diagnostic of immigrants and non-immigrants and subsequently measure the positivity and concreteness of word use for each group. In Study 2, we use the Structural Topic Model (STM) to identify topics among status updates and estimate their relationship to immigration status (immigrant vs non-immigrant) and openness scores of Facebook respondents. Overall, we found words used by non-immigrants to be more concrete and more positive than words used by immigrants; however, once accounted for pair-wise comparisons, these differences resulted as not reliable. Finally, we found significant interactions between openness scores and immigration status among 6 selected topics. Our comprehensive exploration gives new insights on how language use distinguishes people of different immigration status.

Keywords: immigrants, language use, openness, psycholinguistics, stm

3.2 Introduction

Immigrants and non-immigrants show differences throughout some dimensions of the Five Factor model of personality. According to this model, people's personality can be mostly explained in terms of how they position themselves across five dimensions or traits: openness, conscientiousness, extraversion, agreeableness and neuroticism. Big Five traits are commonly used as a standardized instrument to study the psychology of migratory individuals. In fact, several studies in migration psychology have observed population differences in the distribution of personality scores across the Big Five traits. For example, Camperio Ciani et al. (2007) compared Italian populations of islanders and mainlanders and found that immigrants to the islands showed the typical personality profile of mainlanders, i.e., they were more open and extrovert but less conscientious and emotionally stable than islanders. Differences in Big Five traits distribution have also been shown to have some influence on migration intentions (Canache, Hayes, Mondak & Wals, 2013; Fouarge, Ozer & Seegers, 2016; Paulauskaitė et al., 2010). For example, Canache et al. (2013) studied nationally representative samples from 22 countries in the Americas and showed that higher levels of openness and extraversion influence positively the intention to migrate. Big Five traits have also been used to predict whether individuals were going to cross the national borders or not when undertaking a migration (Jokela et al., 2008; 2009). While the literature shows solid consensus about the role of openness regarding migration propensity, there is less clarity about the roles of other Big Five traits. For example, in Chapter 1 we did not find any relation between agreeableness or extraversion and any type of migration considered in this analysis. Moreover, Jokela (2009) found no effect for neuroticism, while both Silventoinen et al. (2008) and Jokela et al. (2008) showed a positive relation between this trait and the tendency to migrate. In sum, openness seems to be a good

predictor of not only underlying differences between immigrants and non-immigrants, but also between lower and higher migration propensity. This suggests that open-minded people who appreciate novelty and challenge are the ones who are more likely to undertake a change of residence.

Another way to tap into psychology and behaviour is observing trends and differences in language use (Coltheart, 1981; Pennebaker et al., 2003; Stone et al., 1996). In fact, language and psychology are strictly connected, as the former represents a reliable tool to manifest thought and emotion in a way that other people can understand (Tausczik & Pennebaker, 2010). Recent studies observed the language used in social media platforms to find what words or topics are more typical for a particular group of people, who are classified based on known variables such as geographical location, gender, age or psychological characteristics. Snefjella et al. (2018) investigated whether national character stereotypes between the US and Canada were reflected in their language use on Twitter. They found, for example, that Canadians tend to use less abbreviations or slang in their usage of verbs, and that Americans tend to use words whose valence is more negative than the valence of words used by Canadians. Other studies found variations in language use across groups of different gender, age and personality profile. Schwartz et al. (2013) analyzed Facebook status updates from approximately 70,000 users to find variations in language use that align with past gender or age studies. They found, for example, that women tend to use more emotional words than men, and that men refer more to objects and use more swear words.

In this study, we use a similar approach to investigate whether differences in personality between immigrants and non-immigrants are reflected in their language use. Using data from Facebook that was collected through the MyPersonality project

(https://www.psychometrics.cam.ac.uk/productsservices/mypersonality), we conduct two analysis to first assess differences in language use between our groups and then introduce another variable, openness, to observe any interaction between personality and immigration status reflected in their word use on Facebook.

3.3 Materials

Participants volunteered to share their demographic information and status updates as part of the MyPersonality project, where they also took surveys that assessed user's personality scores (Kosinski, Stillwell, & Graepel, 2013). Using these data, we selected participants based on their location (i.e., Canada and the USA) and immigration status (i.e., those whose current town and hometown mismatched) to analyze their patterns of language use. Before modelling the data, we engaged in some text processing steps, which consisted of dropping punctuation, removing stop words (e.g., the, at, and) and tokenizing the text (i.e., segmenting the running text into words). Next, we reduced words into their lemmata (e.g., runners > run) and tagged the text for part of speech, morphological and syntactical information using the R package *UDPipe*, version 0.8.2. Syntactic information followed the universal dependency schema (Nivre et al., 2016). Since we are interested in semantic differences in language use, we only kept verbs, nouns and adjectives. After accounting for these adjustments, we further constrained our data to include only words which had a minimum frequency of 50 occurrences and were used by at least 10 different users. Consequently, the final version of our data included 1,134,202 words (N lemmata = 31,225) distributed across 97,311 users. Our use of these secondary data was approved by the McMaster Research Ethics Board (2018-089).

3.4 Study 1

Methods

In the first step of our analysis, we identify words that are statistically over-represented from the immigrant group relative to the non-immigrant group and vice-versa. Based on the difference between the frequency of a given word in two corpora, we adopted the "log odds ratio informative Dirichlet prior" method - or LORIDP - from Monroe et al. (2008). Given two text documents or corpora, the LORIDP method is particularly suitable to detect differences in high frequency words and is commonly used to find words that are statistically over-represented in one document over the other, and vice versa (Monroe, Colaresi, & Quinn, 2008; Jurafsky, Chahuneau, Routledge, & Smith, 2014; Snefjella, Schmitdke, & Kuperman, 2018). Using LORIDP, we estimated the difference between the frequency of word w in the immigrant and non-immigrant sub-corpora via the log–odds–ratio for a given word w, $\delta_w^{(i-j)}$ which is calculated as

$$\delta_w^{(i-j)} = log\left(\frac{y_w^i + \alpha_w}{n^i + \alpha_0 - (y_w^i + \alpha_w)}\right) - log\left(\frac{y_w^j + \alpha_w}{n^j + \alpha_0 - (y_w^j + \alpha_w)}\right)$$

where i is the immigrant corpus, j is the non-immigrant corpus, n_i is the total number of words in the immigrant corpus i, n_j is the total number of words in the non-immigrant corpus j, y_w^i is the frequency count of word w in corpus i, y_w^j is the frequency count of the word w in corpus j, α_0 and α_w are respectively the total number of tokens and the total frequency of the word w in the original Facebook corpus (immigrant and non-immigrant sub-corpora combined). Subsequently, the variance of the log–odds–ratio measure can be estimated as

$$\sigma^2 \left(\delta_w^{(i-j)} \right) \approx \frac{1}{y_w^i + \alpha_w} + \frac{1}{y_w^j + \alpha_w}$$

and finally, we calculated the z-score statistic of the log-odds-ratio of each word as

$$\frac{\delta_w^{(i-j)}}{\sqrt{\sigma^2 \left(\delta_w^{(i-j)} \right)}}.$$

The resulting LORIDP *z*-scores ranged from -89.08 to 43.46, where the word associated with the

lowest negative value (-89.07) represented the highest degree of over-representation among non-

immigrant users and the word associated with the highest positive value (43.46) represented the

highest degree of over-representation among immigrant users. For example, the words *mum,*

exam and *film* resulted with positive scores (z = 43.46, 33.32, 25.42) and therefore they are

significantly over-represented among immigrant users, compared to non-immigrants. On the

other hand, words like *get, today* and *ready* have negative scores (z = -89.08, -65.57, -50.57) and

thus they are more representative of the non-immigrant corpus (see Appendix for list of 50 most

diagnostic words for immigrants and non-immigrants). After applying the Bonferroni correction

for multiple comparisons, we removed all words whose z-scores were included in the non-

significant range of -5.7 ≤ z-score ≤ 5.7.

In the next step, we examine the relationship between how diagnostic a word is of

immigrants or non-immigrants and how concrete and positive that given word is perceived.

Estimates for concreteness of words are taken from Brysbaert, Warriner & Kuperman (2014),

who provide concreteness ratings for 37,058 English words collected from human participants on

the crowdsourcing platform Amazon Mechanical Turk. Combining this dataset with our data yields to 876 words rated for concreteness (N 'immigrant' words = 250; N 'non-immigrant' words = 626). Valence norms are taken from Warriner, Kuperman & Brysbaert (2013), who used similar Internet crowdsourcing methods to collect valence ratings for 13,000 lemmas (i.e., words in their base form). After applying norms from Warriner et al. (2013), we retrieved 391 words rated for positivity (N 'immigrant' words = 141; N 'non-immigrant' words = 250). Concreteness is measured on a 1-5 scale, from abstract to concrete, while valence is measured on a 1-9 scale, from negative to positive. To examine the relationship between how diagnostic of immigration status a word is and how concrete and positive that word is perceived, we aggregated all immigrant and non-immigrant words and divided them into 10 bins, ordered by LORIDP z-score for each of the two groups. Bin 1 consists of the most diagnostic words for both immigrants and non-immigrants; bin 10 consists of the least diagnostic ones. In this way, we were able to compare concreteness and positivity of immigrant and non-immigrant words in each bin, ranked by how diagnostic a group of words in each bin is across the two groups.

Results

Concreteness. Non-immigrant words are generally more concrete than immigrant words, although the difference between the two groups is small (d = 0.24; p < 0.002). Fig 1 reports the concreteness of bins of immigrant and non-immigrant words ranked by LORIDP; Cohen's d and p-values for t-tests within each bin are reported in the right side of the figure. Although some bins present greater difference than others, there is no correlation between how 'concrete' and how diagnostic a bin is of one of the two groups. For example, the difference between words in bin 3 (d = 0.71) is greater than the difference between words in bin 10 (d=0.13), which includes the 10% most diagnostic immigrant and non-immigrant words. However, after applying the

Bonferroni correction for multiple comparisons, p-values associated with each bin turned out to be statistically not reliable ($p > 0.05$). Non-immigrant words are generally more concrete across the overall distribution, with no correlation to how diagnostic of immigrants or non-immigrants words are.

Valence. Overall, non-immigrant words are slightly more positive than immigrant words ($d = 0.26$; $p < 0.001$). Fig 2 presents the valence (positivity) of words across immigrant and non-immigrant bins. Cohen's d and p-values for t-tests within each bin are reported in the right side of the figure. As for concreteness, the Bonferroni correction for multiple comparisons produced a series of non-significant p-values across our bins. Therefore, although a difference in concreteness results from an overall comparison between the two groups, once we divide the words in bins and rank them by LORIDP, no significant difference is reflected at the bin level.
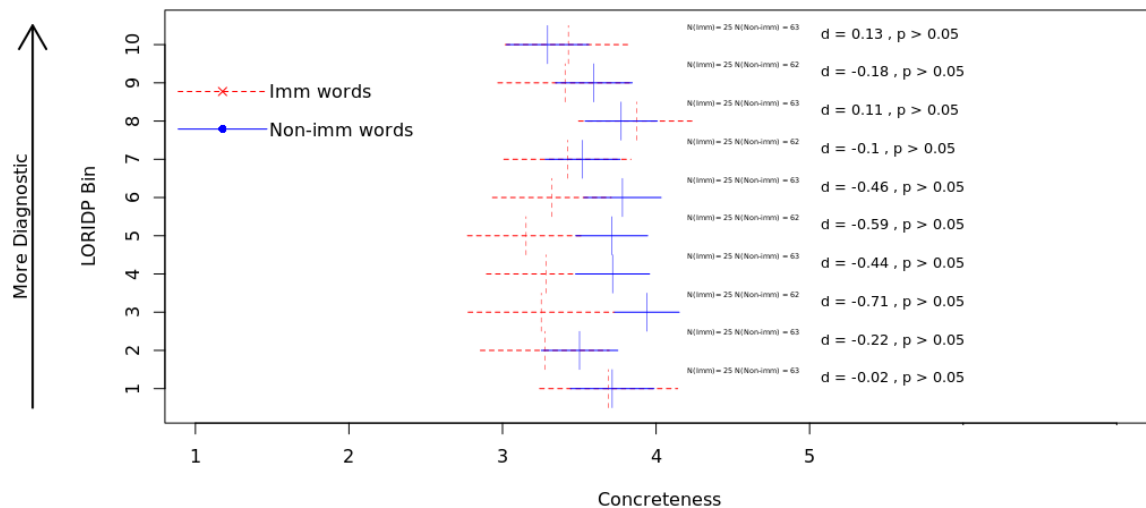

Figure 4. Concreteness of immigrant and non-immigrant words by LORIDP bin. Relative concreteness is shown in red for immigrant words and in blue for non-immigrant words.
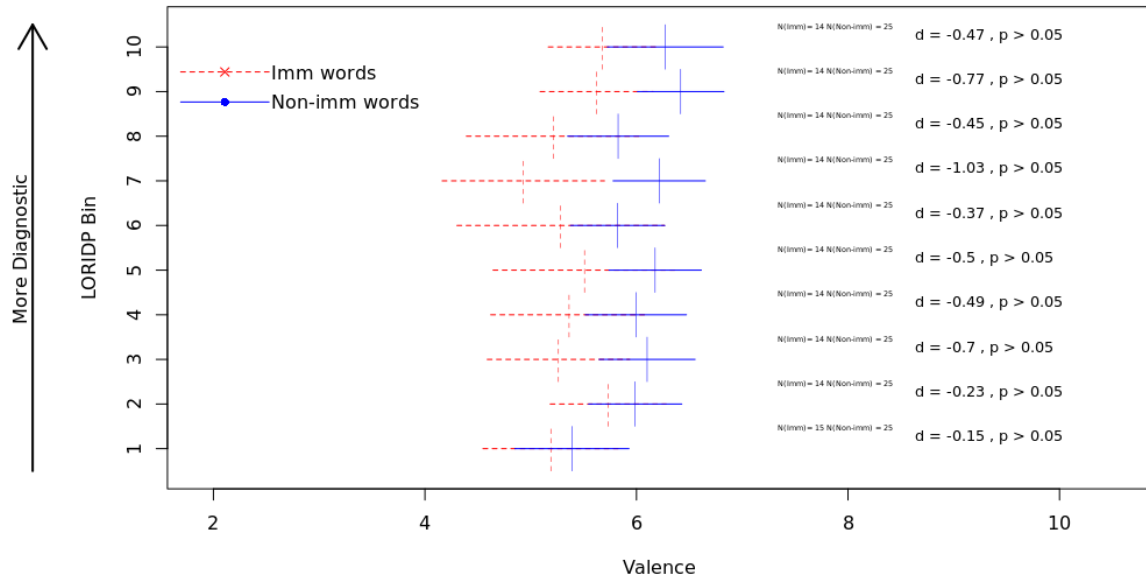
Figure 5. Valence of immigrant and non-immigrant words by LORIDP bin. Relative positivity is shown in red for immigrant words and in blue for non-immigrant words.

3.5 Study 2

Methods

In Study 2, we use the Structural Topic Model (STM) to identify the main topics across Facebook status updates and estimate their relationship with our document-level metadata (in this case, immigration status and openness of Facebook users). Topic models from our Facebook statuses were estimated in the programming environment R (version 3.5.3) and implemented with the *stm* package version 1.3.3.

The STM is classified as a type of open-vocabulary approach: linguistic features (words) in the text are automatically organized in sets of semantically related words (topics). Like other topic models, STM defines a topic as a mixture over words (where each word has a probability of

belonging to a topic) and a document as a mixture over topics (i.e., a single document can be constituted by multiple topics). The STM further allows its users to incorporate additional information (metadata) about each document to the models. This permits not only to discover the most discriminative topics but also to estimate their relationship with the metadata and ultimately use the output of the topic model to test hypotheses on these relationships. Therefore, the inclusion of covariates can contribute to inference and interpretability of the data and can affect both topical prevalence (i.e., how much of a document is associated with a topic) and topical content (i.e., words used within a topic).

In this study, we intend prevalence to be a function of the "immigration status" variable, which is coded as either "Immigrant" or "Non-immigrant," and the variable "openness", coded as personality ratings ranging 0-5. Our first action in Study 2 was to look at collections of words that are associated with topics. There is no right or wrong number of topics that are appropriate to be found in a given corpus (Grimmer and Stewart 2013). To avoid a computationally costly process of identifying an optimal number of topics, we instructed our model to identify 20 topics with the spectral initialization. The output of this estimation is a list of the most characteristic language feature for any topic that discriminates words based our covariates. For each topic, the STM by default prints different types of word profiles, including highest probability words, low probability words and FREX words. Reading these word profiles is helpful for understanding the content of a topic and interpreting its meaning. In particular, FREX words are weighted by their overall frequency and 'exclusivity' to the topic, and therefore represent a good baseline for the STM user who labels a topic with a description (generally one word) that is more meaningful and direct than the word profiles in the topic. From now on in the text, FREX words will be reported next to the respective topic, between parentheses, in italic. For example, we can see that Topic 7

traces discussions of personal and philosophical matters (*knowledge, being, fear, belief, individual, desire, path*) and therefore we labelled this topic as 'Introspection'. Topic 14 discusses everyday questions related to school or sport activities (*homework, practice, soccer, semester, camp, volleyball, senior*) and therefore we decided to label it as "School/Sport".

In the second step, we estimated the relationship between topics and metadata. In this part we observe potential interactions between covariates such that one variable may modulate the effect of the other variable. Specifically, we set our STM to allow for an interaction between openness scores (entered linearly for simplicity) and immigration status for each of the topics of interest. For simplicity, we restricted our analysis to those topics that did not refer to URLs (i.e., Topic 1: *chatter.com, http, www.youtube.com, yearbook.com*), numbers (Topic 10: *1, 2, 5*), emoticons or abbreviations characteristic of the language of instant messages (Topics 17: *xd, dd, xp*, and 18: *wit, jus, tht, ppl, bout*), swear words (Topic 3: *fuck, fucking, shit, bitch, ass*), words whose grouping criterion is unclear (Topic 13: *paste, copy, status, cancer, repost, post, cure*) and linguistic minorities (Topics 5: *une, vous, fait, nous, monde*, 6: *inte, ett, nner, mit, att*, 8, 9, 11, 12, 15 and 19: *porque, gente, vivir, estar*). This restricting process resulted in six topics (Topics 2, 4, 7, 14, 16, 20), for each of which we observed how openness modulated topical prevalence among our two groups. In the following section we plot and discuss metadata/topic relationships resulted from our analysis.

Results

We are interested in looking at how language use differs for immigrants and non-immigrants and how this difference is modulated by openness. To accomplish this, we first specify immigration status as the variable to use for calculating the difference effect. Then, since

topic prevalence for a particular topic is contrasted for two groups, our model includes the expected proportion of a document that belongs to a topic as a first difference type estimate. Figure 3 reports the topical prevalence contrast between immigrants and non-immigrants. Given that our covariate of interest is binary (we are interested in the contrast between immigrants and non-immigrants), Fig 3 includes the change rate in topic proportion shifting from one value to another. We can see how Topics 2 (Sleep/Work), 4 (Family) and 14 (School/Sport) are strongly used by non-immigrants compared to immigrants. Topics 7 (Introspection) and 20 (Happiness) are close to the middle, but the former is immigrant-leaning while the latter is non-immigrant-leaning. Finally, it is worth to notice how our model grouped together non-English words in several languages, finding topics that basically reflect the use of languages other than English on Facebook. Topics 9, 12 and 19 are largely associated with immigrants rather than non-immigrants; in fact, they reflect the use of words that belong primarily to - respectively - Indonesian, British English and Spanish. However, this is just indicative of the fact that whoever is more likely to use words in Dutch or Polish, is also likely to come from the Netherlands or Poland: the contrast with non-immigrants is therefore maximized. This explains part of the why we see discussions about work, school and family contrasting maximally with discussions that contain British English, Dutch or Indonesian words.

We then proceeded in re-estimating the STM to allow for an interaction between openness scores and immigration status. After applying the restrictions discussed in the methods section, we observed how openness modulates topical prevalence among immigrants and non-immigrants in six resulting topics. Figures 4-9 show the graphical display of topical content for topics 2, 4, 7, 14, 16, 20, respectively. P-values for all interactions across the six topics are significant ($p < 0.001$). Given the number of topics evaluated by our model (N = 20), we set the threshold of

significance for topical prevalence at 0.05: all topics discussed less than 5% are not considered particularly informative. We observe that non-immigrants wrote about Topic 2 (Sleep/Work) more than immigrants, but Topic 2 diminishes in salience as openness increases. Conversely, Topic 4 (Family) is largely discussed by immigrants, whereas it does not reach the 5% threshold in non-immigrants. Conversely, Topic 7 (Introspection) is discussed in great measure by both groups; furthermore, it increases in salience as a linear function of openness scores. Topic 14 (School/Sport) is also well discussed among both groups, but more among non-immigrants. Although for non-immigrants topical prevalence decreases as openness increases, topic prevalence for immigrants stays more consistent and therefore is less affected by changes in openness. Topic 16 (Food) is slightly more discussed by non-immigrants, but its topical prevalence increases with a similar rate for both groups. Finally, Topic 20 (Happiness) shows little difference in both topical prevalence and openness for the two groups.
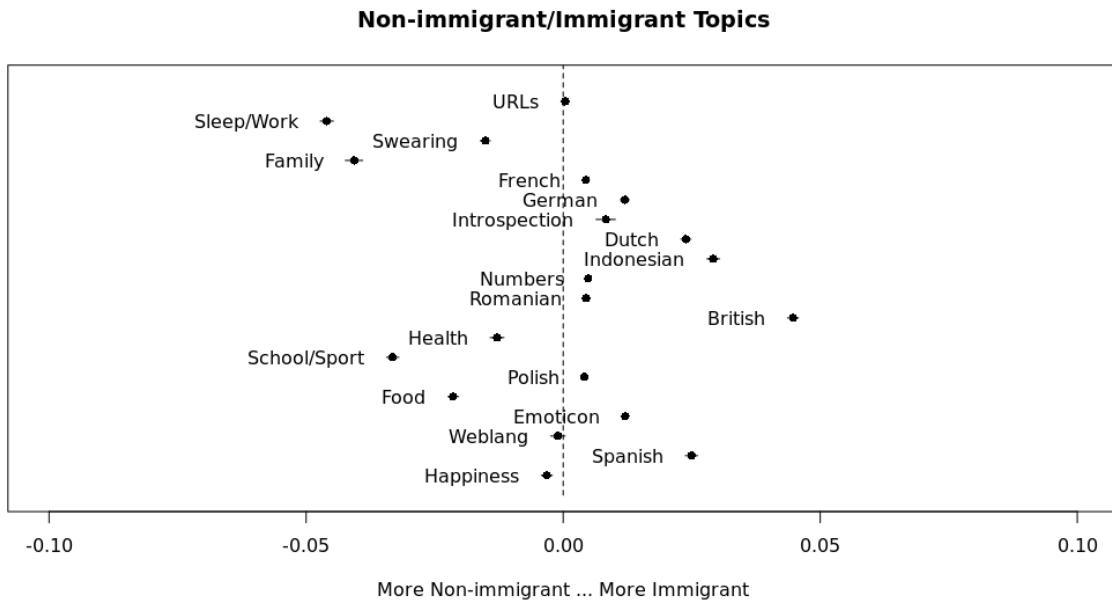
Figure 6. Graphical display of topical prevalence contrast between non-immigrants (left) and immigrants (right).
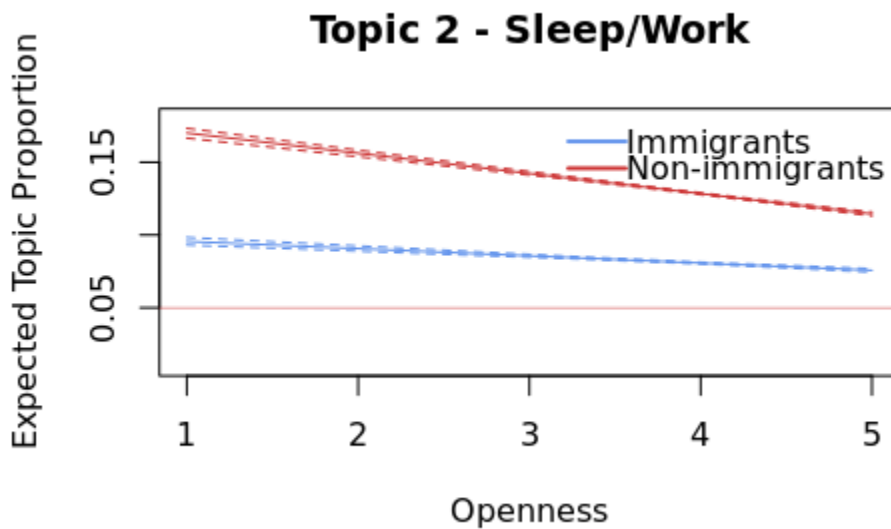

Figure 7. Interaction between openness and immigration status. Topic 2 prevalence is plotted as linear function of openness, holding the immigration status at either "Immigrant" or "Non-immigrant".

Figure 8. Interaction between openness and immigration status. Topic 4 prevalence is plotted as linear function of openness, holding the immigration status at either "Immigrant" or "Non-immigrant".

Figure 9. Interaction between openness and immigration status. Topic 7 prevalence is plotted as linear function of openness, holding the immigration status at either "Immigrant" or "Non-immigrant".
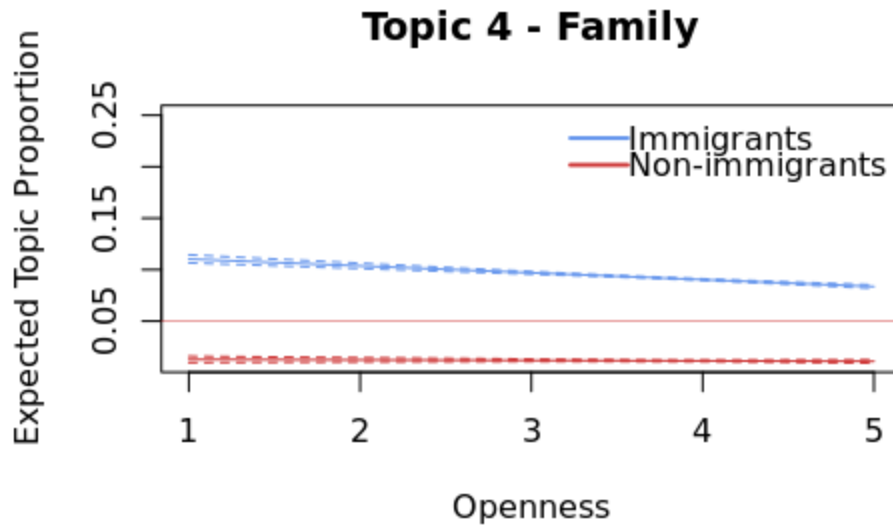
Figure 10. Interaction between openness and immigration status. Topic 14 prevalence is plotted as linear function of openness, holding the immigration status at either "Immigrant" or "Non-immigrant".

Figure 11. Interaction between openness and immigration status. Topic 16 prevalence is plotted as linear function of openness, holding the immigration status at either "Immigrant" or "Non-immigrant".
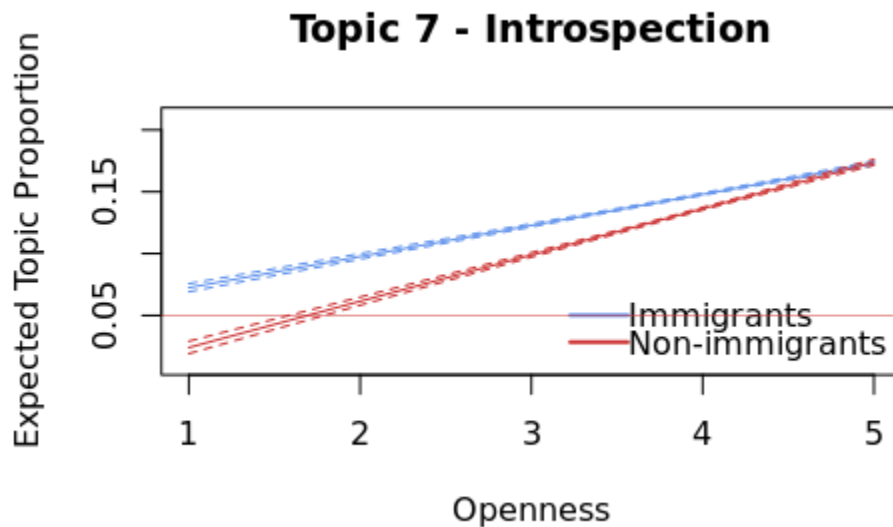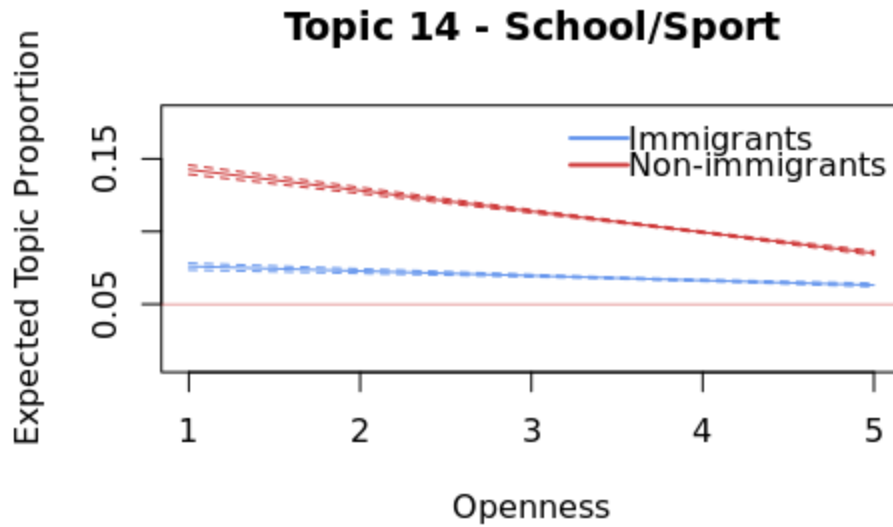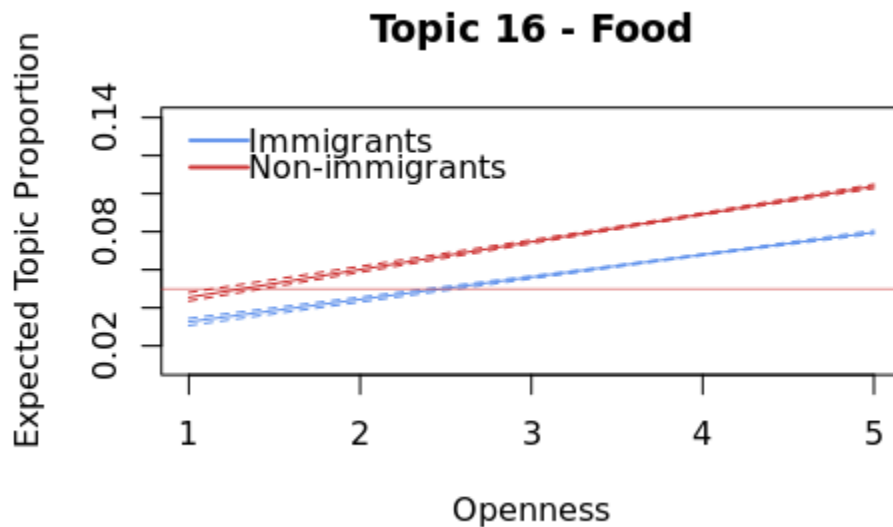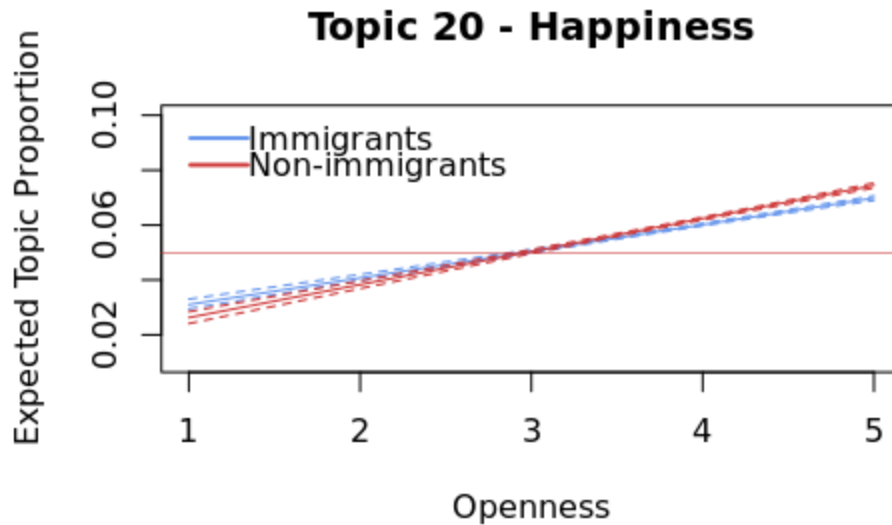
Figure 12.Interaction between openness and immigration status. Topic 20 prevalence is plotted as linear function of openness, holding the immigration status at either "Immigrant" or "Non-immigrant".

3.6 Discussion

Using Facebook data from 97,311 users and 75,665,716 status updates, we conducted two analyses to investigate whether personality differences between immigrants and non-immigrants are reflected in their language use.

In Study 1, we calculated what words are more characteristic of immigrants and non-immigrants and then measured the concreteness and positivity of word use for each group. We found that non-immigrant words are generally more concrete and more positive than immigrant words, although the difference is small between the two groups. Difference in use between abstract and concrete words have been investigated by Kousta et al. (2011), whose theory states that abstract words are grounded in people internal (or affective) state, while concrete words are generally mainly grounded in the contextual, physical environment around us (Kousta et al.,

60

2011; Vigliocco et al., 2013). According to this theory, it is no surprise that words used by non-immigrants resulted as more concrete and positive than words used by non-immigrants. In fact, this difference could be explained by the fact that new immigrants are not as rooted in the social fabric as non-immigrants are, who are naturally more accustomed to use concrete words.

In Study 2, we use the Structural Topic Model (STM) to identify topics among status updates and estimate their relationship to immigration status (immigrant vs non-immigrant) and openness scores of Facebook respondents. We found topical prevalence for some topics to be highly contrasted between the two groups, i.e., some topics were more typical of one of the groups. For example, the "Sleep/Work" and the "Family" topics appear to be the most non-immigrant topics and are highly contrasted to topics such as "British" and "Indonesian". As we anticipated in the results section, this contrast is not suggesting that immigrants don't talk about family matters; rather, it means that our model recognized non-English words and has grouped them accordingly. Therefore, given that non-immigrants generally never use non-English words, the contrast is maximized. We also found openness interacting with immigration status differently based on different topics. For example, both groups talk about the "School/Sport" topic, but for non-immigrants topical prevalence decreases as openness increases, whereas topic prevalence stays more consistent for immigrants; conversely, the "Introspection" topic increases in topical prevalence as a linear function of openness scores for both groups. In other words, openness increases for both groups as they talk more about concepts like belief, thought, desires, choice etc. The same trend is observed for topics 16 ("Food": *butter, peanut, cheese, sauce, cookie, soup*) and 20 ("Happiness": *dance, rehearsal, fabulous, dancing, amazing, sunshine*). Although openness is a good predictor of underlying differences between lower and higher migration propensity, it appears that differences in openness between immigrants and non-

immigrants are minimized for those who talk more about their mental or emotional processes, cheerful events and for some reason, food. In other words, our more diagnostic non-immigrant topics where openness in non-immigrants increases along with topical prevalence are the topics more discussed by non-immigrants who are likely to be open-minded and appreciative of novelty and challenge. This could also suggest that non-immigrants who write about "Introspection", "Food" or "Happiness" are those who are more likely to represent the group of the immigrants-to-be, than those non-immigrants who talk more about topics like "School/Sport" and "Family". As a matter of fact, those who write more about school and family may have families or more established community ties and therefore less inclined to change place of residence (see (Lounsbury et al., 2013). Further, since non-immigrants in first countries don't have the same financial pressure as non-immigrants in poorer countries, personality is likely to play a more important role for wealthy non-immigrants when it comes to migration propensity. As Tabor (2015) suggests, "while financial causes may explain international migration from poorer to richer nations, individual differences are more likely to explain international migration for those already living in relatively rich countries." Therefore, our findings not only align with the notion of openness as a powerful source of individual differences when it comes to migration, but also introduce the idea that certain topics discussed by non-immigrants could be predictive of their desire of novelty and challenge.

There are also some limitations to consider. First, participants were all volunteers, i.e., they all select themselves into one group. This could cause bias in the sampling, since – potentially- features that lead individuals to self-select could cause abnormal conditions in the groups they put themselves into. Further, it is worth noticing that language often depends on context (Eckert, 2008). In relation to this study, this means that the differences and analogies we

report here are limited within the context of social media and Facebook in particular, and maybe they would not be found in context different from this. Although the differences reported are relatively small, our model successfully identified differences between topics that were highly contrasting based on immigration status. The novelty of our study is in investigating language of social media to shed some light on the differences of word use between immigrants and non-immigrants and gives new insights on how certain topics distinguish people of different immigration status.

3.7 References

Brysbaert, M., Warriner, A. B., & Kuperman, V. (2014). Concreteness ratings for 40 thousand generally known English word lemmas. *Behavior research methods*, 46(3), 904-911.

Camperio Ciani, A. S., Capiluppi, C., Veronese, A., & Sartori, G. (2007). The adaptive value of personality differences revealed by small island population dynamics. *European Journal of Personality: Published for the European Association of Personality Psychology*, *21*(1), 3-22.

Canache, D., Hayes, M., Mondak, J. J., & Wals, S. C. (2013). Openness, extraversion and the intention to emigrate. *Journal of Research in Personality*, 47(4), 351-355.

Coltheart, M. (1981). The MRC psycholinguistic database. *The Quarterly Journal of Experimental Psychology Section A,* 33(4), 497-505.

Eckert P (2008) Variation and the indexical field1. *Journal of Sociolinguistics* 12: 453–476.

Fouarge, D., Ozer, M., & Seegers, P. (2016). Personality traits and migration intention: Who bears the cost of migration. Maastricht University, mimeo.

Jokela, M. (2009). Personality predicts migration within and between US states. *Journal of Research in Personality*, 43(1), 79-83.

Jokela, M., Elovainio, M., Kivimäki, M., & Keltikangas-Järvinen, L. (2008). Temperament and migration patterns in Finland. *Psychological Science*, 19(9), 831-837.

Jurafsky, D., Chahuneau, V., Routledge, B. R., & Smith, N. A. (2014). Narrative framing of consumer sentiment in online restaurant reviews. *First Monday*, 19(4).

Kosinski, M., Stillwell, D., & Graepel, T. (2013). Private traits and attributes are predictable from digital records of human behavior. *Proceedings of the National Academy of Sciences*, *110*(15), 5802-5805.

Kousta, S. T., Vigliocco, G., Vinson, D. P., Andrews, M., & Del Campo, E. (2011). The representation of abstract words: why emotion matters. *Journal of Experimental Psychology: General*, *140*(1), 14.

Lounsbury, J. W., Loveland, J. M., & Gibson, L. W. (2003). An investigation of psychological sense of community in relation to big five personality traits. *Journal of Community Psychology, 31*, 531–541.

Monroe, B. L., Colaresi, M. P., & Quinn, K. M. (2008). Fightin'words: Lexical feature selection and evaluation for identifying the content of political conflict. *Political Analysis*, 16(4), 372-403.

Nivre, J., De Marneffe, M. C., Ginter, F., Goldberg, Y., Hajic, J., Manning, C. D., ... & Tsarfaty, R. (2016, May). Universal dependencies v1: A multilingual treebank collection. In *Proceedings of the Tenth International Conference on Language Resources and Evaluation (LREC 2016)* (pp. 1659-1666).

Paulauskaitė, E., Šeibokaitė, L., & Endriulaitienė, A. (2010). Big five personality traits linked with migratory intentions in Lithuanian student sample. International journal of psychology: a biopsychosocial approach, 2010, Vol. 7, p. 41-58.

Pennebaker JW, Mehl MR, Niederhoffer KG (2003) Psychological aspects of natural language use: our words, our selves. *Annual Review of Psychology* 54: 547–77.

Schwartz, H. A., Eichstaedt, J. C., Kern, M. L., Dziurzynski, L., Ramones, S. M., Agrawal, M., ... & Ungar, L. H. (2013). Personality, gender, and age in the language of social media: The open-vocabulary approach. PloS one, 8(9), e73791.

Silventoinen, K., Hammar, N., Hedlund, E., Koskenvuo, M., Rönnemaa, T., & Kaprio, J. (2007). Selective international migration by social position, health behaviour and personality. *European Journal of Public Health,* 18(2), 150-155.

Snefjella, B., Schmidtke, D., & Kuperman, V. (2018). National character stereotypes mirror language use: A study of Canadian and American tweets. PloS one, 13(11), e0206188.

Stone P, Dunphy D, Smith M (1966) The General Inquirer: A Computer Approach to Content Analysis. MIT press.

Tausczik Y, Pennebaker J (2010) The psychological meaning of words: Liwc and computerized text analysis methods. *Journal of Language and Social Psychology* 29: 24–54.

Vigliocco, G., Kousta, S., Vinson, D., Andrews, M., & Del Campo, E. (2013). The

representation of abstract words: What matters? Reply to Paivio's (2013) comment on

Kousta et al.(2011).

Warriner, A. B., Kuperman, V., & Brysbaert, M. (2013). Norms of valence, arousal, and

dominance for 13,915 English lemmas. *Behavior research methods,* 45(4), 1191-1207.

Appendix

Example of 50 most diagnostic words for immigrants and non-immigrants; words are ranked by how diagnostic a word is for each group.

Most diagnostic 'non-immigrant' words	Most diagnostic 'immigrant' words
get	mum
go	bloody
today	exam
work	lovely
day	wee
ready	film
tonight	net
mom	den
school	eh
tomorrow	holiday
great	mate
game	gusto
family	wag
good	nag
fun	brilliant
kid	lad
see	pub
class	cinema
weekend	revision
take	coursework
know	mode
movie	factor
house	assignment
time	pang
hang	jest

guess	match
make	rubbish
homework	revise
night	mat
suck	load
clean	ale
friend	keen
little	mummy
favourite	amigo
bed	fab
excited	mere
morning	bin
need	cum
ugh	advert
week	persona
baby	result
head	tidy
thankful	twat
ass	saber
nap	akin
laundry	cricket
call	main
dinner	gal
start	legend
wonderful	tea

4 General Discussion

The goal of this thesis is to gain insight on the psychology of migratory populations. We focused on three aspects regarding the relationship between personality traits and immigration, and explored our hypotheses in three different studies. Interpretation of the results is provided in more detail in the discussion section of respective chapters. In the following section, we summarize our findings and point out limitations and future directions of the present body of work.

In chapter 1, we asked whether personality traits could predict shorter or longer distances of migration (or no migration at all). Our findings confirmed and refined the hypothesis that different personality traits are associated with different types of migration, from a local level to a more international scale. Specifically, individuals who have changed residence over their lifetime resulted as more open and conscientious than those who never have; our results align with Paulauskaitė et al. (2010) and it applies to migrants who moved within states, between states and also across countries.

In chapter 2, we investigated whether migration had an effect on openness of responders, and how this relationship is modulated by their age and gender. In other words, we asked whether immigrants and non-immigrants differed in their trend of openness as a function of age, and if this difference was reflected when controlling for gender. We found that aging affects both immigrants' and non-immigrants' openness, and that the magnitude of this effect changes at different age windows. Generally, our results show that during early adulthood all immigrants – female and male - are less open than their non-immigrant counterparts. This trend is subject to change with time, as in later adulthood immigrants become more open than non-immigrants.

Finally, in chapter 3 we focused on differences in language use between immigrants and non-immigrants on social media. Our interest was in investigating whether personality differences between immigrants and non-immigrants were reflected in their language use. We found that non-immigrant words are generally more concrete and more positive than immigrant words, although the difference between the two groups is relatively small. We also found what topics are more characteristic of immigrants and non-immigrants users and calculated how openness interacts with immigration status differently, based on different topics. Our results suggest that differences in openness between immigrants and non-immigrants are minimized for those who talk more about certain topics rather than others.

Although limitations that are specific for each study are examined in the discussion section of each chapter, there are some which can be extended to the entire thesis in general. First, our classification of immigrants and non-immigrants is only binary and based on the time when responders shared their information. We considered 'immigrants' the individuals whose original location of residency mismatched with their current (i.e., at the time data was collected) location of residency. Therefore, we ignored whether migrations at the individual level occurred multiple times. Future studies could address this issue by dividing people who migrated only once to North America with those who arrived after having migrated elsewhere and see how they compare. One possible prediction based on results from chapters 1 and 2 could be that differences between 'one-time migrants' vs. 'multiple-times migrants' would level out with age, with 'multiple-times migrants' starting out as scoring higher in openness in the age windows where differences in personality are normally minimized (see Soto et al., 2011). Another possible implementation of this thesis would be in coding the countries from where people moved to North America. As discussed in chapter 2, immigrants to new countries are likely to face

difficulties related to socio-cultural and linguistic factors (Dewaele & Stavans, 2014; Simich et al., 2005; Zhou et al., 2008). One hypothesis worth to be tested could be that people who migrate across cultures where English is not the native language would show larger differences in openness as a reaction to the unfamiliar linguistic context. Finally, one addition to this work could lay in monitoring differences in language over time to draw conclusions about what enhances integration. In fact, no (or minimized) difference in the adoption of abstract and concrete words by both immigrants and non-immigrants could be interpreted as both categories being equally grounded in their contextual, physical environment (see Kousta et al., 2011; Vigliocco et al., 2013). Similarly, longitudinal observations about differences over the same topic could give insights on how immigrants align with the linguistic and social fabric of their new country.

In conclusion, the present thesis has both corroborated existing findings and added novel ones to the knowledge about psychology of migratory populations. My work has confirmed that the trait openness to experience is strictly connected to underlying motivations behind one's distance of migration and migration propensity. It also demonstrated that migration is a determining factor that influences the development of openness across different age groups and genders. Finally, it considers the language of social media, proposing that contrasts in language use between immigrants and non-immigrants are reflective of deeper dissimilarities - and in some cases, analogies – between the two groups. These results have implications for theories about the connection between immigration and personality, suggesting that individual differences in psychological configuration play a determinant role in both the phases preceding and following the migration. Therefore, personality traits and language use should be considered

in any attempt of uncovering underlying motivations, desires, differences and analogies behind migratory and non-migratory populations.

References

Barrick, M. R., & Mount, M. K. (1991). The big five personality dimensions and job performance: a meta-analysis. *Personnel psychology*, *44*(1), 1-26.

Berry, J. W. (2001). A psychology of immigration. *Journal of social issues*, *57*(3), 615-631.

Bodvarsson, Ö. B., & Van den Berg, H. (2013). *The economics of immigration*. New York, NY: Springer New York.

Boneva, B. S., & Frieze, I. H. (2001). Toward a concept of a migrant personality. *Journal of Social Issues, 57*(3), 477-491. doi: 10.1111/0022-4537.00224

Borjas, G. J. (2014). *Immigration economics*. Harvard University Press.

Boyle, P. (2009). Migration. In K. Rob & T. Nigel (Eds.), *International Encyclopedia of Human Geography* (pp. 96-107). Oxford: Elsevier.

Brettell, C. B. (2013). Anthropology of migration. *The Encyclopedia of Global Human Migration*.

Brettell, C. B., & Hollifield, J. F. (2013). Theorizing migration in anthropology: The social construction of networks, identities, communities, and globalscapes. In *Migration theory* (pp. 121-168). Routledge.

Camperio Ciani, A. S., Capiluppi, C., Veronese, A., & Sartori, G. (2007). The adaptive value of personality differences revealed by small island population dynamics. *European Journal of Personality: Published for the European Association of Personality Psychology*, *21*(1), 3-22.

Dewaele, J. M., & Stavans, A. (2014). The effect of immigration, acculturation and

multicompetence on personality profiles of Israeli multilinguals. *International Journal of Bilingualism*, *18*(3), 203-221.

Fawcett, J. T. (1985). Migration psychology: New behavioral models. *Population and Environment*, *8*(1-2), 5-14.

Freeman, G. P. (1995). Modes of immigration politics in liberal democratic states. *International migration review*, *29*(4), 881-902.

Goldberg, L. R. (1990). An alternative" description of personality": the big-five factor structure. *Journal of personality and social psychology*, *59*(6), 1216.

Hollifield, J. F. (2000). The politics of international migration. *Migration theory: Talking across disciplines*, 137-85.

Jokela, M. (2009). Personality predicts migration within and between US states. *Journal of Research in Personality*, *43*(1), 79-83.

Jokela, M., Elovainio, M., Kivimäki, M., & Keltikangas-Järvinen, L. (2008). Temperament and migration patterns in Finland. *Psychological Science*, *19*(9), 831-837.

Otto, K., & Dalbert, C. (2012). Individual differences in job-related relocation readiness: The impact of personality dispositions and social orientations. *Career Development International, 17*(2), 168-186. doi: 10.1108/13620431211225340

Ożańska-Ponikwia, K., & Dewaele, J. M. (2012). Personality and L2 use: The advantage of being openminded and self-confident in an immigration context. *Eurosla Yearbook*, *12*(1), 112-134.

Silventoinen, K., Hammar, N., Hedlund, E., Koskenvuo, M., Rönnemaa, T., & Kaprio, J. (2007). Selective international migration by social position, health behaviour and personality. *European Journal of Public Health*, *18*(2), 150-155.

Simich, L., Beiser, M., Stewart, M., & Mwakarimba, E. (2005). Providing social support for immigrants and refugees in Canada: Challenges and directions. *Journal of Immigrant and Minority Health*, *7*(4), 259-268.

Soto, C. J., John, O. P., Gosling, S. D., & Potter, J. (2011). Age differences in personality traits from 10 to 65: Big Five domains and facets in a large cross-sectional sample. *Journal of personality and social psychology*, *100*(2), 330.

Stark, O., & Taylor, J. E. (1989). Relative deprivation and international migration. *Demography, 26*(1), 1-14.

Tabor, A. S. (2010). A framework for voluntary migration: Understanding modern British migration to New Zealand. Masters Thesis, Victoria University of Wellington, Wellington, New Zealand.

Tabor, A. S., & Milfont, T. L. (2011). Migration change model: Exploring the process of migration on a psychological level. *International Journal of Intercultural Relations, 35* (6), 818-832. doi: 10.1016/j.ijintrel.2010.11.013

Verhoeven, L., & Vermeer, A. (2002). Communicative competence and personality dimensions in first and second language learners. *Applied Psycholinguistics*, *23*(3), 361-374.

Zhou, Y., Jindal-Snape, D., Topping, K., & Todman, J. (2008). Theoretical models of culture

    shock and adaptation in international students in higher education. *Studies in higher*

    education, *33*(1), 63-75.