**Measuring Immigrant Integration**

October 24, 2016

Lucila Figueroa

Rachel Gillum

Jens Hainmueller

Dominik Hangartner

David D. Laitin

Duncan Lawrence

Abstract

This paper proposes a standard measure of immigrant integration – i.e. the degree to which immigrants have the knowledge and the capacity to achieve success in their host society – that permits the comparison of immigrant communities over time and across contexts. To justify our measure, we first show the costs for a research community when every study relies on its own specification of what constitutes successful integration. We then adumbrate the criteria for a successful measure. Once set, we outline six dimensions of integration—psychological, economic, political, social, linguistic and navigational – each with a set of survey questions. With these questions, we run pretests to determine the degree to which our questions meet our criteria for a good measure, and to reduce our tool to twelve key questions that could be incorporated in all studies of integration at low cost. We report on the data we have so far collected and the issues they raise. However imperfect, we foresee substantial payoffs for scientific progress of community “buy in” for our measure.

**Note Well: this is a working lab document of a project in early development. Comments welcome.**

**1. Introduction:**

 The premise of this paper is that the research community would profit from a standard measure of integration—the degree to which immigrants have the knowledge and the capacity to achieve success in their host society—a measure that researchers who study immigration into the advanced industrial countries of the global north lack. Without such a measure, each study relies on its own specification of what constitutes successful integration, substantially reducing the possibility of comparison across countries and over time. Cumulative knowledge about rates of integration success is consequently lacking.

Justifications for the current heterogeneity of definitions and proxies are usually based on recognition that integration as a concept is “essentially contested” (Gaillie 1955-56) or too complex to be reined in by a single metric. This, however, is equally true in the measurement of wealth, but scholarly agreement on either GDP or HDI – indices that have substantial construct validity – permit well-conceived causal analyses showing *inter alia* the effects of policy on growth. One useful model is the “Kessler” index for mental health; its “big six” questions have been shown to have construct validity and are broadly used in a wide variety of studies, thereby permitting comparability. Our goal is to facilitate such research on the integration of immigrants.

 We cannot overemphasize the critical need for such research. The mass of refugees that have been entering Europe from Afghanistan, Syria, and North Africa have made this issue of global importance. Meanwhile, the flow of refugees from embattled zones in Central America into the United States, combined with the long-standing and large-scale immigration from Mexico to the U.S., presents similar challenges. We observe cities, countries, and regional organizations devising responses to both newer refugee population flows and historical migration patterns that are *ad hoc* and varied. Now is the time to learn the implications of the various policies and procedures for immigrant welfare and societal peace. Alas, empirical studies currently being conducted use different measures haphazardly, hampering comparability and accumulation of knowledge. Without standard measures of the outcomes we care about, it is impossible to ascertain “best practices” in integrating immigrants into the host societies in Europe and North America.

 Our contribution is a proposed pragmatic survey instrument that fulfills a set of goals. It captures several of the multifarious dimensions of integration—psychological, economic, political, social, linguistic and navigational — and permits theoretically motivated scales for each dimension that demonstrate high levels of correlation among the survey questions behind each dimension. This allows us to pare down the questions on each dimension to allow for a sparse sub-set of questions that best capture each dimension. If any dimension of integration is being studied as a causal variable, the index score for that dimension would be available. If integration is to be studied as an outcome variable, the expectation would be a future index (comparable to the “big 6” in the Kessler index) combining the dimensions. The goal is to capture, each with a single question, the two principal components on each dimension. This implies a module of questions that we call the Immigration Policy Lab, or IPL-12.

 We emphasize here that our measure does not claim to be the only, best, or perfect measure of integration. The key is that it is useful and has the capacity (through research community buy-in) to generate cumulative knowledge. Immigration research needs such a measure to make progress. The key is not so much that the measure is perfect but, having met several criteria outlined below, that it establishes a standard.

 We believe the returns for both causal inference and policy relevance will be high if there is academic “buy in” for a common integration module. We would be able to quantify the impact of policies on integration (e.g. by comparing similar host communities with similar integration rates but relying on different incorporation policies). We could also design micro-studies of randomized local treatments (for example, the encouragement of permanent legal residents to apply for citizenship), and analyze their long-term impacts on integration. If performed in a variety of places with different immigrant groups (and again, using a common measure of integration), we could determine whether policy success differs depending on the group or the host society. In all these proposed studies, we could estimate policy effects with a degree of confidence that we do not have in studies currently available that lack a common metric.

 To foreshadow our conceptual stance, we distinguish assimilation (the adoption of the cultural practices of the dominant group by newcomers) from integration. Our practical definition of integration is *the degree to which immigrants have the knowledge and the capacity to achieve success in their host society.* Knowledge entails aspects such as fluency in the national language and ability to navigate through public and private institutions in order to get bank accounts, admission into hospitals, drivers’ licenses, legal status renewals, and voting registration. Capacity refers to the resources immigrants have to make investments in their (and their children’s) futures and the degree to which the host society lowers the entry barriers to immigrant inclusion in economic, social and political life.

**2. Literature review**

 The literature on the integration of immigrants into the advanced industrial countries in the global north is immense, and cannot be thoroughly reviewed in the course of a single article. There are three interlinked literatures that we build on. First is an extensive set of surveys, most of them collected by state agencies throughout Europe and North America, that seek to measure the degree to which immigrant populations are integrating into their societies. We list the datasets that we have consulted and from which we have appropriated variables for our proposed modules in the Appendix. The major strengths of these datasets are that they reveal the questions that are important to policy makers and statistical bureaus throughout the advanced industrial world, and that they provide tables of descriptive statistics allowing us to see which questions have greatest inferential yield (at least for the country where the data were collected). The major weakness of these datasets is that they tend to be oriented to the specific issues of the countries in which they were administered, thereby reducing their use in comparative analysis.

 The second literature we have consulted seeks to bring conceptual clarity to the core concepts in use (integration, incorporation, and assimilation). It then employs currently available national datasets along with surveys attached to particular research projects in order to analyze cross-country and over-time trends in the immigrant experience. Exemplary of this genre is *Outsiders no more? Models of Immigrant Political Incorporation* (2013) a volume that represents where the academic literature now stands in regard to the sociology of immigrant incorporation. A core goal of the book, according to the editors’ introduction, is to develop a standard approach for “measuring incorporation.” They recount the substantial difficulties of finding a measure since there is disagreement as to what constitutes someone as an immigrant; disagreement as to what precisely immigrants are incorporating into; and disagreement about how much political involvement is necessary for immigrants before they achieve incorporation. In their own review of the literature on incorporation, they mourn “one may find almost as many usages … as there are authors” (2013, 7).

 To their credit, the editors take a stand on a definition of political incorporation, *viz*. “having the capacity for sustained claims making about the allocation of symbolic or material public goods” (2013, 16). As X. Briggs emphasizes in his chapter (2013, 321-342), the editors’ definition of political incorporation advances the concepts of membership and of the capacity for influence in the political arena. On this latter point, he argues “such capacity helps guard against political domination by the few and the subversion of core values, and it tends – over time and space – to be associated with more equitable access to ‘opportunity’” (Briggs 2013, 325).

 Indeed we rely on this notion of capacity in our definition of integration. Despite this useful conceptual contribution, however, the editors offer no scales of “degrees of incorporation” that can be measured across cases or time. Instead, the editors celebrate the pluralistic approach to defining incorporation, their outcome variable. This is hardly helpful if we want to accumulate knowledge. From the collection of essays, many of which are insightful in their own terms, we therefore have almost no way of knowing whether group *a* in country *b* is more incorporated than group *a’* in country *b*, or group *a* in country *c*. To be blunt, however insightful the accumulated studies in this volume, there can be no scientific progress or useful policy prescriptions without a common standard for success.

 As a result of the lack of a measure of their outcome, the present state of knowledge would be as if we had extensive discussions of the wealth of nations without an agreed upon standard of GDP or of HDI. No economist believes these measures are perfect proxies for the vague concept of “wealth”, but without agreement in the scientific community of a measure for wealth, no theory can be tested. We can make a similar point in regard to the psychological literature on mental health, captured in the “Kessler Index”.

 The goal of this paper is to build a community of researchers who will rely on a common measurement of what we call integration that can be incorporated in all surveys conducted in this field along with a range of other questions based on the particular research question being asked. This measure would have several dimensions: psychological, political, economic, social, linguistic and navigational integration. Each dimension would contain a short module of questions. Having a common outcome measure (“*Y*”) across a range of studies would, in our judgment, allow for the accumulation of findings. In sum, progress in this field is contingent on community acceptance of how to measure *Y*; theory advances as we examine the range of causal factors (*X*’s) that may matter for variance in *Y*, with the *abc*’s used as controls.

 The third literature we consulted, one in which substantial progress in measurement has been made, is the comprehensive National Academy of Sciences report on the integration process (National Academies 2015). Though focused on the American experience, this NAS report has clear lessons for future work on integration throughout the world. The report provides a broad definition of integration: “The [NAS] panel defines integration as the process by which members of immigrant groups and host societies come to resemble one another (Brown and Bean, 2006).”[[1]](#footnote-1) In making this definition applicable to empirical research, the panel distinguishes among fourteen concepts related to integration success though it recognizes that it often is conflating two separate dimensions of change: integration and well-being. The first asks whether immigrants and the native-born become more like one another; the second asks whether immigrants are better or worse off over time.

 Integration success as analyzed in the NAS report includes those fourteen indicators across five broad dimensions. First, there is economic integration. This includes variables on income, living above the poverty line, and occupational distribution (i.e. the degree to which immigrants move into job sectors with the same distribution as the native-born population). Second there is social integration. This includes variables on civic engagement including local volunteerism, marriage (degree to which immigrants marry outside their cultural group), family patterns (and how they differ from those of native born), and spatial integration (defined by the extent to which residence patterns are ethnically segregated). Third, there are a set of variables that measure the attainment of human capital. They include language ability, health, and educational attainment. Fourth, two variables measure a degree of psychological integration. They include the degree to which respondents report “feeling American” and feeling safe (reporting of degree to which they feel subject to crime). Fifth, the report includes measures of political integration. These refer to legal status and their political participation as voters, as candidates, and as elected representatives.

 The authors of the NAS study recognize issues in conceptualizing their outcome variable as assimilation. First is the problem of coming to terms with the enormous diversity in all host societies, and the mean values of such things as economic success or security from crime for members of the host society obscure enormous in-group variation. The NAS panel was not blind to this, and were concerned that mean values for health tend to be higher for some groups of immigrants than for native born. Should we infer from this, they wonder, that integration would require those immigrants to suffer from health declines? Similarly, suppose certain immigrant communities have much higher rates of civic engagement than the modal citizen. Should that indicate a lower level of integration?

 Nonetheless, the NAS’s fourteen components have a high degree of construct validity in capturing our intuitions about integration. We seek to take advantage of the conceptual distinctions that guided the NAS research (that indeed reflects mostly the sociological literature on integration) but does so in a way that is more efficient in capturing the differing dimensions of integration across a wide variety of political and social contexts (i.e. with a minimal number of survey questions) and doing so without slipping into the confusion of assimilation with that of integration.

**3. Our Approach: Theory and Methodology**

 Although our definition deviates somewhat from that of the NAS study, we draw from that study (and from the national surveys that we reviewed) a set of six dimensions to measure integration success: economic, social, psychological, political, linguistic, and navigational.

 In devising questions for each of these dimensions, we were guided by six criteria. First, the questions in our module need to reflect construct validity, i.e. the degree to which the question is actually measuring integration, and not something else. To assure ourselves that questions meet this standard, we sought questions from existing surveys that had distributions of responses more or less commensurate with the range of answers to similar questions. In other words, we discounted questions that yielded unusual distributions of responses compared to related questions as probably capturing something different from what was intended. Another check comes from our pilot surveys. If we find that respondents who have been in their host country longer systematically reveal themselves less integrated in their host society (other things equal) we would investigate whether the question was measuring something different from our intentions.

 Second, the questions required clear directionality. We see this clearly in our measures of linguistic and navigational integration, where each level of host country language facility or each step in navigating the health care system of the host country signals higher levels of integration. However, our pilots have made us wary as to whether one measure of social integration – being members of clubs with folks from different home countries – has the same feature of directionality. It may be that well-integrated second- or third-generation immigrants think of home country as the same as their host country, and report that their club members (all from the host country) are from the same home country as theirs. In this case, the most integrated in our sample would appear, by our measure, as integration failures on the social dimension. This question would then lack directionality.

 Third, the questions should not presuppose that cultural repertoires of the society in which immigrants were integrating are to be emulated. To do so would be to conflate “integration” with “assimilation”. Doing so is a problem for two reasons. First, the notion of achieving success in a host country is conceptually different from adoption of the cultural practices of the dominant group in the host society. Second, and relatedly, burdening the integration measure with aspects of cultural adoption excludes the possibility of asking whether strategies of assimilation are associated with (or even cause) high degrees of integration. The more narrowly the concept of integration is conceived, the easier it is to study the relationship between it and related processes.

 But to avoid conflating these two related concepts turned out to be a challenging criterion. We rejected the notion popularized by Huntington in *Who Are We?* (2004) in claiming that resembling Protestants of European origin was the benchmark for integration. Moreover, why should we think that baseball (rather than soccer) or apple pie (rather than tacos) represent America’s core culture? Indeed Schildkraut (2005, 2011) reports that many rooted Americans contest values that Huntington sees as core to American culture. Even in Europe, where some hold to a myth of a “true” cultural core (e.g., Lebovics 1992*)*, defining that core is no simple matter. The attempt by President Sarkozy of France in using social media to determine what constitutes a French person to serve as a model for integration was a nonstarter as few could agree to a common notion of who or what was truly French.

 To preview our solution, we ask immigrants how they conceive of their home (or ancestral) country identities, and then ask if they have social and other forms of associations with native born citizens of their host country from different home country backgrounds without privileging any ethnic, racial or religious heritage as core. This approach boils down to whether immigrant families develop network ties with host country native born citizens from different home country backgrounds. We believe this solution captures our everyday notion of integration.

 Fourth, our questions need to translate well into different national and local environments. Asking about drivers licenses for immigrant communities living in world cities has a different meaning from those living in suburbs. We thus needed to avoid questions that would reveal differences due to local conditions rather than integration success.

 Fifth, the questions need to be adjustable to differing modes of administration – they need to be applicable for phone surveys, door-to-door surveys, and surveys done through mobile technologies. Asking respondents to report their interpretation of a picture, for example, would work on a tablet but not on a phone interview.

 Sixth and finally, our questions need to yield variation across responses. One of our key concerns in examining the descriptive statistics in national surveys was whether the answers across immigrant respondents had a normal distribution. The greater a question was able to distinguish among different levels of integration, the more useful it would be in future analysis.

**4. The Questionnaire**

In this section we provide an overview of our initial pilot questions associated with each of the dimensions that we derived from the literature. We administered the survey in the United States using the online survey software, Qualtrics. For the purposes of this study, respondents’ “host country” refers to the United States and its citizens as Americans. This of course would be altered for administration in other countries.

The questionnaire began with a set of standard demographic questions concerning gender, age, country of birth, date of arrival in host country, type of visa upon first arrival, citizenship status, education (and in which country was it provided), marriage/relationship status, nationality of spouse/partner, number and given names of children, aspirations for their children’s education, zip code, and national origin of parents and grandparents. It then seeks a benchmark for the ancestral country with which the respondent most identifies, what we refer to as respondent’s “pipe country”.[[2]](#footnote-2) None of the answers to these questions go into our proposed index; rather they are the background information that is essential for any interpretive exercise.

 The questionnaire then moves on to the dimensions of integration, the first being *psychological*. We first ask, providing a five-point Likert scale for answers, “Please tell us how much you agree or disagree with the following statement: ‘I feel a close connection with the United States.’” We then ask respondents their degree of generalized trust, their trust of Americans, and then their trust in people of their pipe country. (Since the generalized trust question comes first, we can anchor responses to host/pipe country with the generalized trust measure. Our notion is that the difference between trust in host and pipe country indicates degrees of psychological identification with one as opposed to the other). Still on the psychological dimension, we ask a series of questions, most with Likert-scaled answers: “In general, how welcoming are most Americans towards people like you?”; “When Americans meet you for the first time, do they typically think of you as a fellow American?”; “How important is it to you that your children are perceived as Americans” (alternatively, a hypothetical version of the question appears for those who do not have children);[[3]](#footnote-3) “Are you planning on living in the U.S. forever, or do you plan on leaving the U.S. permanently at some point”; “In what country would you like to be buried or have your ashes laid to rest?” We then ask respondents to divide among 100 points the degree to which they identify with their host and their pipe country.

 Still on psychological integration, we ask how often respondents felt they were treated with less respect than other people; whether people acted as if they were not smart; whether people seemed afraid of them; whether they were ever threatened or harassed; and whether they thought they were a member of a group that is treated less fairly than other groups in American society. Of course there are many reasons why someone may feel that their group is treated less fairly and we therefore also ask respondents to identify whether their group is treated unfairly due to their race/ethnicity, nationality, religion, status as an immigrant, language, age, gender, sexual orientation, income, education level, or other.

 The second dimension is *economic* integration. We ask about respondents’ housing situation (rent or own; amount of space; number of household members living there); their links to the financial system (whether they have a bank account in their own name in the U.S.); their employment (employment status, their current job category and their last job category in their home country if they are immigrants, type of contract; number of hours; income; whether salaried or self-employed; job security; degree to which job category is commensurate with their skills); the degree to which they face job discrimination due to aspects of their identities; their reliance on other household members for income support; and their ability to meet everyday expenses.

 The third dimension is *political* integration. Here we underline two considerations in the development of this module. First, in differentiating assimilation from integration, we hold that while assimilation may entail becoming disengaged or apathetic about politics, engagement in the political process makes one integrated. This implies (correctly in our judgment) that the modal American is not well integrated into the political process. Second, we see political integration as not only entailing engagement, but also knowledge of host country political institutions. On engagement, we ask if respondent votes (in various types of elections); whether s/he has participated in a range of political activities (signed a petition, boycotted a product, communicated with a government official, worked in a political party/action group/other political organization, worn or displayed a campaign badge/sticker, taken part in a lawful public demonstration, collected signatures); whether s/he discusses politics with others; and whether she or he has an ideological and partisan identification in the host country political system. On knowledge, we ask respondents to place the two major parties on a left-right ideological scale, where we consider respondents who place Democrats to the left of Republicans as more knowledgeable, regardless of how left or right respondents think the parties are overall. We then ask if they know what party controls the executive and the legislative branches, and the eligibility age for voting in the United States. We also ask several questions derived from a standard battery commonly used in political science research to measure efficacy; questions about impressions of the police (whether police officers treat “people like you” with respect, whether police officers make respondents feel safe, and whether respondents would call the police if they needed help); and how often respondent listens to political news about their host country.

 The survey then moves on to a measure of *social* integration. Here we ask if any family member has married a native born citizen of the host country whose origin is different pipe country.[[4]](#footnote-4) Similarly, we ask whether respondent thought that those Americans from different pipe countries and/or racial/ethnic backgrounds would be open to marrying in their community. We then ask about membership in a variety of social organizations, each time questioning the degree to which respondent interacts with U.S.-born American citizens from different pipe country backgrounds. In the same vein, we ask about the people respondents have depended on in the recent past, such as acquaintances, friends, relatives, and professionals (e.g. ministers and social workers), for help with everyday favors such as getting rides, borrowing a little money, babysitting, or getting a job, again with a focus on interactions with people from different racial/ethnic/national origin backgrounds. Finally, for social integration we ask about co-workers, neighbors, classmates of their children, and friends on social media, each time asking the degree to which they are US-born Americans from different backgrounds.

 Fifth, we develop measures of *linguistic* integration. Linguistic integration is the component of cultural integration that meets the required criteria of our instrument. Namely, non-linguistic measures of cultural integration fail to meet the criteria that questions should have a clear directionality, should avoid conflating integration with assimilation, and should translate well into different national and local environments. For example, while a commitment to individualism is a major component of American culture (de la Garza, Falcon, and Garcia 1996; Lipset 1990; Schildkraut 2005), measuring whether immigrants in the U.S. are individualistic does not translate well into different national environments where individualism is not a core component of host country values. Additionally, while the U.S. is a very religious country, and while many believe the U.S. is a Christian country, for many other Americans the freedom to practice any religion, or no religion at all, is a core component of the United States. Therefore, there is no clear directionality to measuring religion as a component of cultural integration. Additionally, suppose one were to posit that since most Americans are Christian, immigrants are more integrated if they are also Christians. Such an assertion fails to meet the criterion that integration is not synonymous with assimilation. Finally, focusing on religion as a component of American culture does not translate well into more secular countries. Language, on the other hand, is an important component of all cultures, and a *sine qua non* for successful advance in any host society.

 Linguistic integration encompasses two components: usage and proficiency. In terms of usage, we ask about the language respondents use on social media, in speaking with their children, with their own primary caregiver when they were young, their closest friend in the host country (and the country of origin of that friend), their coworkers, and their supervisors at work. This leads to a series of questions about the degree of difficulty respondents face in communicating (in the local *lingua franca*) with doctors, teachers, and police. Afterwards, explicitly focusing on proficiency, we ask respondents about their skills in reading, listening, writing and speaking the dominant language of their host country (or the linguistically distinct region of their host country).[[5]](#footnote-5)

 Finally, our sixth dimension is that of *navigational integration*, seeking to measure the degree of success immigrants achieve in their dealings with both informal and formal institutions of their host country, enabling them to function in it efficiently. We ask respondents the degree of difficulty they face in opening a bank account, in making a doctor's appointment, in obtaining a driver's license, in using public transportation, in enrolling their children in a school, in getting vaccines/immunizations, in finding adequate housing, in finding a job, and in coordinating childcare.

 Each of these six dimensions of integration–psychological, social, economic, political, linguistic, and navigational–that we have adumbrated need to be indexed. After we present descriptives from two pilot surveys relying on our questionnaire, we propose a schema for doing so. Since we find that each of our dimensions has at least two principal components, we hope to identify two key questions on each dimension which captures the two major factors embedded in each dimension and maximizes (a) construct validity; (b) correlation with the other questions on that dimension; (c) variance across responses on that question; and (d) robustness on the above criteria across countries and over time. These would become the core questions for our “IPL-12” integration measure.

 As a final question in our present survey, we ask: “Thinking about your life today, how integrated do you feel in the United States? Please indicate your level of integration on the scale below, where 0 means you feel not integrated at all and 10 means you feel fully integrated.” Hopefully, this will correlate strongly with our “IPL-12” index; yet we expect the “IPL-12” will come closer to the underlying factor of integration and also allow for a degree of variance in the score on each dimension.

**5. Analysis of the Survey Sample**

 We drew from two different samples to pilot our integration instrument. First, we obtained a sample through Stanford University’s Research Experience Program (REP). The REP is a joint program between Stanford University, Foothill Community College and Mission College, to provide human subjects to Stanford researchers. Students who are enrolled in introductory social science classes at these two community colleges participate in REP research for course credit. Community college students took the survey online, where we used Qualtrics software to administer the survey. There are 187 respondents in the REP sample.

 We also purchased a sample of 1,025 respondents through Qualtrics from Survey Sampling International LLC (SSI). We implemented quotas on the collection of the SSI sample, requiring that a minimum of 400 respondents be immigrants, at least 400 be second generation Americans (U.S.-born with at least one foreign-born parent), and a minimum of 200 people be third generation Americans (U.S.-born with U.S.-born parents and at least one foreign-born grandparent). We include second and third generation individuals because it is not only the integration of immigrants themselves that we care about, but also about the ability of immigrant groups to integrate over time and over successive generations. All respondents in the SSI sampling were living in the United States.

Table 1: Generational Status

|  |  |  |  |
| --- | --- | --- | --- |
|  | Total | SSI | REP |
| Generation | N | % | N | % | N | % |
| 1st | 461 | 38.13 | 409 | 39.94 | 52 | 28.11 |
| 2nd  | 493 | 40.78 | 410 | 40.04 | 83 | 44.86 |
| 3rd  | 221 | 18.28 | 205 | 20.02 | 16 | 8.65 |
| 4th­ +  | 34 | 2.81 | 0 | 0.00 | 34 | 18.38 |

 We pool the two datasets together, and end up with a total of 1,209 respondents, as displayed in Table 1. Thirty-eight percent of the respondents are foreign born, 41% are second generation, and 18% are third generation. Also included in the REP sample are 34 fourth+ generation respondents, whom we also refer to as rooted natives. Rooted natives, or fourth+ generation, are people without a recent immigration history in their family. Fourth+ generation respondents, their parents, and all of their grandparents were born in the United States.

 Both the REP and SSI samples have substantial variation in the generational status of respondents, thereby covering the gamut of integration levels to some extent. However, a weakness of both samples is that the surveys were administered only in English. Therefore, respondents in both samples skew towards slightly higher levels of integration, as they are proficient enough in English to successfully get through the survey. Moreover, REP respondents are enrolled in community college programs, meaning they skew towards being more highly educated. We are planning to account for these weaknesses in future data collection.

 Figure 1 presents the national origin groups represented in the data, displaying the pipe countries that comprise 1% or more of the sample.[[6]](#footnote-6) As a reminder, the pipe country refers to the ancestral country for U.S.-born individuals and the country of birth for immigrants. Mexico is the most well-represented pipe country, followed by China and Germany.[[7]](#footnote-7) While Figure 1 shows the two samples pooled together, there are differences between the REP and SSI samples. Asian countries make up four of the top five REP pipe countries—China (15% of the REP sample), Vietnam (13%), the Philippines (12%), Mexico (8%) and India (5%).[[8]](#footnote-8) Conversely, the top five pipe countries in the SSI sample are all western European countries or part of the Americas—Mexico (7% of SSI sample), Germany (7%), Canada (5%), the United Kingdom (5%), and Italy (5%).

 Figure 2 is analogous to Figure 1, but only includes immigrants whose country of origin comprises at least 1% of immigrants in the sample. Whereas Mexico was the most well-represented country of ancestry for U.S.-born respondents (comprising 8.6%), as displayed in Figure 1, it drops down to the sixth spot in terms of country of origin for immigrant respondents (comprising 5.13%). China and the Philippines continue to be near the top, and India, which took spot number nine in Figure 1, shoots up to the third place spot in Figure 2. While there is some shifting around among the countries between the two graphs, most countries that appear near the top in one graph appear near the top of the other. It is somewhat surprising that Mexico does not comprise a larger share of the sample, as Mexican immigrants comprise a large share of all immigrants in the United States (28%[[9]](#footnote-9)). On the other hand, among the top ten largest immigrant groups in the United States, all but two appear in Figures 1 and 2 (and most appear near the top of the figures).[[10]](#footnote-10)





 Table 2 presents selected demographic information for respondents. While women greatly outnumber men in the REP sample, the SSI sample was more evenly divided, helping even out the overall balance in the combined samples. Nonetheless, there is a difference of almost 15 percentage points between the genders, with more women in the data.

The majority of respondents are U.S.-born citizens, almost a quarter are naturalized citizens, and 8% of respondents are permanent residents (green card holders). Respondents who chose “other status” were given the opportunity to write their immigration status in an open-ended text box, and four people indicated (or implied) that they entered the U.S. without documentation.

The respondents are, overall, well-educated, with almost half having completed undergraduate education (and just under 20% have advanced degrees). A quarter went to college, but did not graduate, and 15% went no further than high school.

Most respondents are white (59%), although within the REP, most are Asian, reflecting the high percentage of Asians living in the greater San Francisco Bay area where the sample was collected. Asian respondents are overrepresented as the second largest group in the combined samples (26%), followed by Latinos (16%).

 Finally, the REP sample is much younger than the SSI sample. In the combined sample, on average, respondents are in their early forties (with a median age of 38).[[11]](#footnote-11)

Table 2: Respondent Demographics

|  |  |  |  |
| --- | --- | --- | --- |
|  | Total | SSI | REP |
| Gender | N | % | N | % | N | % |
|  Men | 514 | 42.41 | 459 | 44.78 | 55 | 29.41 |
|  Women  | 690 | 56.93 | 559 | 54.54 | 131 | 70.05 |
|  Other  | 8 | 0.66 | 7 | 0.68 | 1 | 0.53 |
| Documentation Status[[12]](#footnote-12) |  |  |  |  |  |  |
|  Natural born citizen | 757 | 62.56 | 624 | 60.88 | 133 | 71.89 |
|  Naturalized citizen | 298 | 24.63 | 276 | 26.93 | 22 | 11.89 |
|  Work visa | 17 | 1.40 | 14 | 1.37 | 3 | 1.62 |
|  Family visa | 16 | 1.32 | 15 | 1.46 | 1 | 0.54 |
|  Green card | 97 | 8.02 | 85 | 8.29 | 12 | 6.49 |
|  Refugee | 1 | 0.08 | 1 | 0.10 | 10 | 0.00 |
|  Student visa | 15 | 1.24 | 4 | 0.39 | 11 | 5.95 |
|  Tourist visa | 2 | 0.17 | 2 | 0.20 | 0 | 0.00 |
|  Other status | 7 | 0.58 | 4 | 0.39 | 3 | 1.62 |
| Educational Attainment |  |  |  |  |  |  |
|  No degree | 30 | 2.48 | 30 | 2.93 | 0 | 0.00 |
|  High school grad | 151 | 12.48 | 117 | 12.48 | 34 | 18.38 |
|  Some college | 305 | 25.21 | 210 | 25.21 | 95 | 51.35 |
|  Vocational school | 65 | 5.37 | 56 | 5.37 | 9 | 4.86 |
|  Associate’s degree | 75 | 6.20 | 59 | 6.20 | 16 | 8.65 |
|  Bachelor’s  | 346 | 28.60 | 319 | 28.60 | 27 | 14.59 |
|  Postgraduate degree[[13]](#footnote-13) | 236 | 19.50 | 232 | 19.17 | 4 | 2.16 |
| Race/Ethnicity\* |  |  |  |  |  |  |
|  White | 713 | 58.97 | 650 | 63.41 | 63 | 34.24 |
|  Black | 65 | 5.38 | 61 | 5.95 | 4 | 2.17 |
|  Latino/Hispanic | 198 | 16.38 | 171 | 16.68 | 27 | 14.67 |
|  Asian | 313 | 25.89 | 206 | 20.10 | 107 | 58.15 |
|  Native American  | 13 | 1.08 | 11 | 1.07 | 2 | 1.09 |
|  Arab/Mid. Eastern | 11 | 0.91 | 8 | 0.78 | 3 | 1.63 |
|  Pacific Islander | 8 | 0.66 | 5 | 0.49 | 3 | 1.63 |
| Age  | Mean | Median | Mean | Median | Mean | Median |
|  | 42.79 | 38 | 46.04 | 43 | 24.83 | 21 |

\*Respondents were able to mark as many racial/ethnic categories as they identify with. Therefore, there is overlap among the various categorizations.

**6. Indices of Integration**

The next step is to present results from the long version of the existing instrument to find which questions within each module best match the criteria of maximizing (a) construct validity; (b) correlation with the other questions on that dimension; (c) variance across responses on that question; and (d) robustness on the above criteria across countries and over time. After presenting results of the survey, the next section will specify which questions best qualify for the larger set of 24 questions (IPL-24) and for the pared down IPL-12.

**6.1 Construction of integration indices**

For each set of questions that measure the same integration dimension—psychological, economic, political, social, linguistic, and navigational—we used Principal Component Analysis (PCA) to combine the outcome measures into a single scale. This type of averaging across responses that measure the same construct reduces measurement error in survey responses. We use a polychoric variant of PCA that uses linear combinations of the polychoric correlation matrix of the items, rather than the items themselves, to extract the principal components. The advantage of polychoric PCA over regular PCA for most variants of factor analysis is that it can seamlessly accommodate both categorical and continuous data. This is an important feature given the very different scales of our original integration questions, from binary to categorical to (quasi-)continuous.

Figure 3 displays, separately for each integration dimension, the Eigenvalues of the first ten components. For the linguistic and navigational integration index, the first principal component explains more than 69% (linguistic) and 66% (navigational), respectively, of the total variance, and the explanatory power drops sharply afterwards and flattens for the higher-numbered components. The substantive implication of these results is that for the social, linguistic and navigational dimension a single component is sufficient to measure successful integration. A radically different pattern emerges for the indices measuring psychological, economic, and political integration. For these dimensions, the first three (psychological and economic index) or four (political index) components obtain Eigenvalues higher than 1, and the first component typically explains, with 32% (psychological), 33% (economic), and 28% (navigational), only a relatively small fraction of the total variance. For these integration dimensions, a single component is not able to capture most of the relevant variation in the index. With regard to the social integration index, most of the variance is explained by the first two components (41% and 19%, respectively, of the variance). While most of the social integration questions load positively on the first dimension (e.g. intermarriage, help, and those measuring social interactions with US born natives), it seems to be the case that membership in various clubs and associations mostly load on the second dimension. This pattern might also explain some of the counterintuitive findings for the social integration index discussed further below.

Figure 3: Eigenvalues of first ten components of the integration dimension-specific PCA

|  |  |  |
| --- | --- | --- |
| Psychological integration index../../IM/results/psy_screeplot.pdf | Economic integration index../../IM/results/eco_screeplot.pdf | Political integration index../../IM/results/pol_screeplot.pdf |
| Social integration index../../IM/results/soc_screeplot.pdf | Linguistic integration index../../IM/results/lin_screeplot.pdf | Navigational integration index../../IM/results/nav_screeplot.pdf |

Notes: The figure shows the Eigenvalues of the first ten components for each of the six integration dimensions.

**6.2 Face validity: Correlation of integration indices with residency**

As a first test of the face validity of each of the six indices, we inspect their correlation with the respondents’ residency status and residency period. Consistent with most previous work on immigrant integration, we make the uncontroversial assumption that, on average, immigrants who have been living in the country for longer and have permanent residency or citizenship are better integrated. Hence, we interpret a positive correlation between the integration indices and longer residency and more permanent status/citizenship as partial evidence for the validity of our measures.

Figure 4 displays the means and corresponding 95% confidence intervals for each of the six indices stratified by residency status: citizen since birth, naturalized citizen, green card, and other visas. The estimates for the psychological, economic, political, linguistic and navigational indices follow the expected pattern: U.S. citizens, since birth or naturalized, are better integrated than permanent residents who are better integrated than holders of temporary visas. The social integration index is a partial exception here, with green card holders less integrated than born and naturalized citizens but also less integrated then temporary visa holders.

Figure 4: Average integration by residency status

|  |  |  |
| --- | --- | --- |
| Psychological integration index../../IM/results/psy_status.pdf | Economic integration index../../IM/results/eco_status.pdf | Political integration index../../IM/results/pol_status.pdf |
| Social integration index../../IM/results/soc_status.pdf | Linguistic integration index../../IM/results/lin_status.pdf | Navigational integration index../../IM/results/nav_status.pdf |

Notes: This figure shows means (red dot) and 95% confidence intervals (blue lines) of each integration index by residency status.

Figure 5 presents the analogous results by immigrants’ residency period. For the psychological, economic, political, linguistic and navigational indices, the results are in line with the expected pattern: the longer respondents have been living in the US, the higher their integration status. The social integration index is, again, a clear outlier here where we find no significant, and if anything a negative, correlation between the length of the residency period and the integration status.

Figure 5: Average integration by residency period.

|  |  |  |
| --- | --- | --- |
| Psychological integration index../../IM/results/psy_yearus.pdf | Economic integration index../../IM/results/eco_yearus.pdf | Political integration index../../IM/results/pol_yearus.pdf |
| Social integration index../../IM/results/soc_yearus.pdf | Linguistic integration index../../IM/results/lin_yearus.pdf | Navigational integration index../../IM/results/nav_yearus.pdf |

 Note: This figure shows means (red dot) and 95% confidence intervals (blue lines) of each integration index by residency period.

**6.3 What we learned from our pilot**

First, we learned that the questions associated with all dimensions had construct or face validity, as seen in the Tables showing that years in the host country and legal status correlated with our integration measures.

Second, we learned that each of our dimensions were largely multi-factorial, and this led us to seek the best two questions for each dimension – that is an IPL-12 rather than an IPL-6, our original goal.

Third, several issues reduced our ability to rely on our pilot for choosing the IPL-12. We tested in only one host country (the US) and could not determine whether there would be face validity in other environments. Second, we ran our tests with groups that were already fluent in English, and therefore could not see if our questions on other dimensions of integration would distinguish low and high linguistic integration. Relatedly, our sample was highly educated, with similar issues in seeking measures that would maximally distinguish different levels of integration. Finally, we under-surveyed immigrants born in Mexico, so we did not even have a representative sample of US immigrants.

Fourth, we had many questions in which a substantial number or respondents could not provide an answer, e.g. those asking about babysitting or language used with children, or job classification. And questions with dichotomous answers failed to sufficiently differentiate levels of integration success across respondents. We needed for an index with only twelve questions to be sure all respondents would be able to provide responses to all questions and to give answers that would provide maximum discrimination.

Finally, we reconfigured the “social integration” scale that lacked “directionality”. We found, for example, that early arrivals identifying as foreigners were embedded in home country networks; but later on, if they re-identify as native, they are also largely embedded in home country networks.

**6.4 The IPL-24 and IPL-12: Which questions capture most of the integration index?**

Based on what we learned from our pilots, we created an IPL-24 that met our new criteria, and singled out twelve questions (IPL-12) – two from each dimension, that met our intuitions about the two principal components. They are listed below. Where responses are in Likert or Likert-like scales, we omit the choices.

**Psychological Integration:**

 ***IPL 12***

*Psych 1*: Please tell us how much you agree or disagree with the following statement: "I feel a close connection with the United States."

*Psych 2*: How permanent do you see your primary residence in the US?

 ***IPL 24***

*Psych 3*: We'd like to know how much you trust people. Generally speaking, how often can you trust Americans?

*Psych 4*: Thinking about your life today, how integrated do you feel in the United States?

**Economic Integration:**

 ***IPL 12***

*Econ 1:* Which of these descriptions best applies to what you have been doing for the last four weeks? Please select only one.

* In paid work (or away temporarily) (employee, self-employed, working for your family business) (1)
* In school, (not paid for by employer) even if on vacation (2)
* Unemployed and actively looking for a job (3)
* Unemployed and not actively looking for a job (4)
* Permanently sick or disabled (5)
* Retired (6)
* In community service or military service (7)
* Doing unpaid housework, looking after children or other persons (8)
* Other (9) \_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_

*Econ 2:* What is the gross annual income from all sources for your household?

***IPL 24***

*Econ 3:* Now we'd like to ask you about various expenses that your household can or cannot afford - this gives us an indication of how living standards are changing. Please indicate whether your household can or cannot afford to pay an unexpected, but necessary, expense of

|  |  |  |
| --- | --- | --- |
|  | Yes, can afford (1) | No, cannot afford (2) |
| $400 (1) |  |  |
| $1,000 (2) |  |  |
| $10,000 (3) |  |  |
| $50,000 (4) |  |  |

*Econ 4:* Think about the jobs you would like to have or qualify for in the future here in the US. Which best describes your expectations about these jobs?

1. In this country, I feel I can get any job that I qualify for

2. In this country, there are only a few jobs that are blocked for people of my background

3. In this country, it is possible for people of my background to get excellent jobs that we qualify for, but we need to be far better than those with deeper roots in the country

4. In this country, people from my background face many limitations in getting jobs that we qualify for

5. In this country, people from my background are mostly limited to a very few job categories

**Social Integration**

 ***IPL 12***

*Soc 1:* If you think about the people you regularly interact with socially, how many of them are Americans?

*Soc 2:* In the last two months, how often were you treated with less respect than other people based on how Americans viewed your background?

 ***IPL 24***

*Soc* 3: People sometimes participate in different kinds of groups or associations. For each group listed below, do you participate in a group activity at least once per week, at least once per month, at least once per year, belong to the group but do not actively participate, or do not belong to nor participate in the group.

 A trade union, business, or professional organization;

 A church or other religious organization

; A sports, leisure, or cultural group;

 Another voluntary organization

*Soc 3a:* If you think about these groups you are participating in, are all of them Americans, most of them, about half of them, a few of them or none of them?

*Soc 4:* Many people help each other with everyday favors such as getting rides, borrowing a little money, or babysitting. In the past 12 months, how often have you relied on an American for such an everyday favor?

**Political Integration:**

 ***IPL 12***

Please tell me how much you agree or disagree with each of the statements below:

*Pol 1:* "I have a pretty good understanding of the important political issues facing the United States."

*Pol 2:* "People like me don't have any say about what the government does in the United States."

 ***IPL 24***

*Pol 3:* How often do you typically discuss major political issues facing the United States with others?

*Pol 4a:* In politics people often talk about a left-right dimension. Can you please tell me which of the two sentences is true? [For all host countries, choose the leading two parties in the legislature]

The Republican Party is considered to be left of the Democratic Party

The Republican Party is considered to be right of the Democratic Party

*Pol 4b:* To which of the following parties does the current President of the United States belong?

**Linguistic Integration**

 ***IPL 12***

Communicating in English has many components, like reading, listening, writing, and speaking skills.  For the next few questions, please evaluate your own skills in English.

Can you do the following when **reading**, **speaking**, **listening**, or **writing** English very well, moderately well, a little well, or not well at all?  Please mark one answer for each row.

*Ling 1:* I can **read** and understand the main points in simple newspaper articles on familiar subjects.

*Ling 2:* In a conversation, I can **speak** about familiar topics and express personal opinions.

 **IPL-24**

*Ling 3:* I can **write** letters about my experiences, feelings, and about events.

*Ling 4:* I can **listen** to and understand the main points in radio or TV programs about familiar subjects.

**Navigational Integration**

 ***IPL 12***

In this country, how difficult or easy would it be for you to do each of the following?

 *Nav 1:* Find adequate housing

 *Nav 2:* Make a doctor's appointment

 **IPL-24**

 *Nav 3:* Use public transportation

 *Nav 4:* Open a bank account

**7. Conclusion**

 This paper proposes a standard measure of immigrant integration – i.e. the degree to which immigrants have the knowledge and the capacity to achieve success in their host society – that permits the comparison of immigrant communities over time and across contexts. To justify our measure, we first showed the costs for a research community when every study relies on its own specification of what constitutes successful integration. We then adumbrated the criteria for a successful measure. Once set, we outlined six dimensions of integration—psychological, economic, political, social, linguistic and navigational – each with a set of survey questions. With these questions, we ran pilots to determine the degree to which our questions met our criteria for a good measure, and to reduce our tool to twelve key questions that could be incorporated in all studies of integration at low cost. We reported on the data we have so far collected and the issues they raised. However imperfect, we foresee substantial payoffs for scientific progress of community “buy in” for our future measure.

 Before finalizing our integration measurement tool, we are planning to run our updated survey with samples that are more diverse than we have up till now secured. We plan on running run an updated pilot for the Swiss Immigrant Survey which is a stratified sample of all Swiss immigrants. We are also seeking a sample from students in English as a Second Language (TESL) courses (where we have newly arrived immigrants as well as Green Card holders seeking to upgrade their language skills for purposes of the citizenship interview) in New York State’s Opportunity Centers run by the Office of New Americans. With further pretesting and analysis as to whether our preferred questions are meeting the six criteria set in section 3, we hope to provide students of immigrant integration a tool permitting the accumulation of policy-relevant knowledge. With our finalized tool, we will then be seeking “buy in” from experts in this field, both in political science and sociology.

**References** (partial listing)

Briggs, Xavier de Souza (2013). In Hochschild, J., Chattopadhyay, J., Gay, C., and Jones-Correa, M. (2013) *Outsiders No More?: Models of Immigrant Political Incorporation* (New York: Oxford University Press)

Brown, S.K. and Bean, F.D. (2006) Assimilation Models, Old and New: Explaining a Long-Term Process. *Migration Policy Institute*.

de la Garza, R. O. , Falcon, A., & Garcia, F C. (1996). “Will the real Americans please stand up: A comparison of Anglo and Mexican American support for core American values” *American Journal of Political Science*.

Gallie, W.B. (1955-56) *Essentially Contested Concepts*. Proceedings of the Aristotelian Society.

Hochschild, J., Chattopadhyay, J., Gay, C., and Jones-Correa, M. (2013) *Outsiders No More?: Models of Immigrant Political Incorporation* (New York: Oxford University Press)

Huntington, S. (2004) *Who are We? The Challenges to America’s National Identity* (London: The Free Press)

Lebovics, H. (1992) *True France* (Ithaca: Cornell University Press)

Lipset, S.M. (1990) *Continental Divide: The Values and Institutions of the United States and Canada* (New York: Routledge)

Migration Policy Institute tabulation of data from the U.S. Census Bureau's 2010 and 2014 American Community Surveys, and 2000 Decennial Census. Data for 1960 to 1990 are from Campbell J. Gibson and Emily Lennon, "Historical Census Statistics on the Foreign-Born Population of the United States: 1850 to 1990" (Working Paper No. 29, U.S. Census Bureau, Washington, DC, 1999). <http://www.migrationpolicy.org/programs/data-hub/charts/largest-immigrant-groups-over-time>

National Academies of Sciences, Engineering, and Medicine (2015). *The Integration of Immigrants into American Society. Panel on the Integration of Immigrants into American Society*, M.C. Waters and M.G. Pineau, eds. Committee on Population, Division of Behavioral and Social Sciences and Education. Washington, DC: The National Academies Press. doi: 10.17226/21746.

Schildkraut, D. (2005) Press ‘One’ for English: Language Policy, Public Opinion and American Identity (Princeton University Press)

Schildkraut, D. (2011) Americanism in the Twenty-First Century: Public Opinion in the Age of Immigration (Cambridge University Press).

Appendix

|  |
| --- |
| Surveys/Studies Reviewed |
| Abrams, Ando and Hinke (1998) "Psychological Attachment to the Group: Cross-Cultural Differences in Organizational Identification and Subjective Norms as Predictors of Workers' Turnover Intentions." |
| American National Election Studies |
| Bagnall (2015) "Global Identity in Multicultural and International Educational Context" |
| Bergami and Bagozzi (2000) "Self-categorization, affective commitment and group self-esteem as distinct aspects of social identity in the organization" |
| Caselli (2012) Measuring the Integration of Immigrants: Critical Notes from an Italian Experience |
| Children of Immigrants Longitudinal Study (CILS) |
| European Union minorities and discrimination survey (EU-MIDIS) |
| General Social Survey (GSS) |
| Immigrant Citizens Survey (ICS) |
| International Social Survey Programme's (ISSP) National Identity Survey  |
| Kuo and Margalit (2012) "Measuring Individual Identity: Experimental Evidence" |
| Latino National Survey (LNS) |
| Longitudinal Survey of Immigrants to Australia (LSIA) |
| Longitudinal Survey of Immigrants to Canada (LSIC) |
| Longitudinal Survey of the Integration of First-time Arrivals (ELIPA) |
| Mapping Communities Survey (Wong)  |
| Mays and Cochran (2001) "Mental Health Correlates of Perceived Discrimination Among Lesbian, Gay, and Bisexual Adults in the United States" |
| Mental Health Correlates of Perceived Discrimination Among Lesbian, Gay, and Bisexual Adults in the United States |
| National Asian American Survey (NAAS) |
| National Immigrant Survey (NIS) |
| National Survey of Midlife Development in the United States |
| National UnDACAmented Research Project  |
| New York City Department of Consumer Affairs Financial Needs Survey |
| Pew Muslim American Survey |
| Pew National Survey of Latinos |
| Prentice and Miller (1994) "Asymmetries in Attachments to Groups and to their Members: Distinguishing between Common-Identity and Common-Bond Groups" |
| Spain National Immigrant Survey |
| UK Fourth National Survey of Ethnic Minorities (FNSEM) |
| World Values Survey (WVS) |

1. The report is not overly concerned with getting a definition that clearly distinguishes integration from related concepts such as assimilation; on this, they refer to work by Alba and Nee (2003). But, as we argue below, immigrants need not emulate the cultural styles of the dominant groups in society for successful and meaningful lives. Thus, as will become clear, our definition of integration is closer to Kymlicka (2012) who sees success as the “full opportunity for immigrants and their offspring to achieve a wide range of social, economic, and political mobility in their host society.” In other words, ‘integration’ implies that all barriers to full participation in society have been removed.

 [↑](#footnote-ref-1)
2. If all reported family is from the same home country, our Qualtrics tool assumes that this is the respondent’s “pipe country”. However, if between parents or among grandparents there are different home countries, we ask respondent to choose the one s/he identifies with most. We are testing whether a single question – viz., which foreign country, home to at least some of your ancestors, do you identify with the most? – will be sufficient to establish a valid pipe country. [↑](#footnote-ref-2)
3. Hypothetical version of the question for respondents who do not have children: “If you had children, how important would it be to you that your children were perceived as Americans?” [↑](#footnote-ref-3)
4. We initially worried that asking respondents about social interactions with people from the pipe country versus with people beyond the pipe country, because doing so would code some people as more integrated than they truly are. For example, a person who lives in an ethnic enclave in the U.S. is less socially integrated than a person who lives in a more heterogeneous neighborhood (or a homogenous neighborhood, so long as the residents of the homogenous neighborhood tend to be of different racial, ethnic and national backgrounds than the person in question). In the U.S., there are Latino ethnic enclaves that are made up of people from a variety of Latin American countries. If we only asked respondents the degree to which they interact with people who are not from their pipe country, we may code a Mexican immigrant who lives in a Latino ethnic enclave and therefore regularly interacts with Central Americans, for example, as more integrated than would meet our criterion of construct validity. From that viewpoint, by instead asking about interactions with people from different racial, ethnic, and national backgrounds than the respondent, we could better account for the existing racial and ethnic patterns of residential and social segregation. However, we run into the problem (discussed earlier) that well integrated subjects that consider themselves “Americans” could report that they belong to a homogeneous network of “Americans” that would be coded as non-integrated, again violating construct validity. We continue to experiment with our measures of social integration. [↑](#footnote-ref-4)
5. We consulted existing literature on the best practices for measuring language proficiency to build our measures. First, there is debate about the best methods for efficiently testing language proficiency, and second there is the question of whether people are capable of accurately self-assessing their language abilities (Alderson 2005; Bachman 1985; Blanche 1988; Brantmeier et al. 2011; Brown 1983; Brysbaert 2013; Delgado et al. 1999; Fotos 1991; Gaillard and Tremblay 2005; Gogin, Estrada, and Villarrea 1994; Hanania and Shikhani 1986; Heilenman 1983; Oller 1973; Ross 1998; Snodgrass and Vanderwart 1980; Tremblay 2011; Yamashita 1996). Nonetheless, there does seem to be a general consensus that self-assessments using “can-do tasks” – self-assessments that are specific and focused – do a good-job of measuring language proficiency (Brantmeier et al. 2011; LeBlanc and Painchaud 1985; Pierce et al. 1993). “Can-do” tasks ask respondents how well they can do various tasks. For instance, respondents are asked to evaluate whether they “can understand familiar words and very simple sentences, for example, on posters” very well, moderately well, a little well, or not well at all. They are also asked whether they can “read quickly through long and complex texts, locating the details I need to find” very well, moderately well, etc. The former can-do task is simple, the latter is more difficult and serves as a measure of high language proficiency. By asking respondents a series of questions about reading, writing, speaking, and listening abilities, each time asking about how well they can do specific tasks of varying levels of difficulty, one receives an accurate picture of respondents’ language abilities in a relatively efficient manner. This approach is relatively efficient because widely accepted language proficiency tests, like the Test of English as a Foreign Language (TOEFL), are too lengthy to include in an integration instrument. [↑](#footnote-ref-5)
6. The following countries comprise fewer than 1% of the sample and are not included in Figure 1: Brazil, Greece, Spain, Austria, Hungary, Romania, Nigeria, Colombia, the Netherlands, Pakistan, Thailand, Afghanistan, Trinidad and Tobago, Switzerland, Armenia, Australia, Barbados, Belgium, Czech Republic, Denmark, Egypt, El Salvador, Guatemala, Nicaragua, Norway, Peru, Sweden, American Samoa, Andorra, Belarus, Finland, Israel, Lebanon, Lithuania, New Zealand, Aruba, Chile, Iceland, South Africa, U.S. Virgin Islands, Angola, Bosnia and Herzegovina, Bulgaria, East Germany, Ecuador, Ghana, Guyana, Haiti, Honduras, Indonesia, Iran, Iraq, Jordan, Malaysia, Panama, Syria, Tonga, U.S. Minor Outlying Islands, Union of Soviet Socialist Republics, United Arab Emirates, Zimbabwe, Algeria, Antigua and Barbuda, Bangladesh, Bermuda, Bolivia, Cambodia, Canton and Enderbury Islands, Congo-Kinshasa, Croatia, Djibouti, Georgia, Grenada, Guam, Laos, Latvia, Liberia, Malta, Moldova, Mongolia, Nepal, North Korea, Panama Canal Zone, Saint Lucia, Saint Martin, Saint Vincent and the Grenadines, Senegal, Serbia, Singapore, Slovenia, Sri Lanka, Sudan, Togo, Uruguay, and “other country.” [↑](#footnote-ref-6)
7. China includes the special administrative regions of Hong Kong and Macau. [↑](#footnote-ref-7)
8. China is the most well-represented REP country of ancestry (15% of the sample), followed by Vietnam (13%), the Philippines (12%), Mexico (8%), and India (5%). All other countries constitute fewer than 5% of the REP sample. [↑](#footnote-ref-8)
9. http://www.migrationpolicy.org/programs/data-hub/charts/largest-immigrant-groups-over-time [↑](#footnote-ref-9)
10. The only two countries that are among the top ten largest immigrant groups in the U.S. that do not comprise at least 1% of the sample are El Salvador and Guatemala. The remaining eight countries are (in descending order) Mexico, India, China, the Philippines, Vietnam, Cuba, South Korea, and the Dominican Republic (Migration Policy Institute). [↑](#footnote-ref-10)
11. Numerous REP respondents are senior citizens, with six of them claiming to be over one hundred years old (as well as three in the Qualtrics data). Despite concern that these respondents are untruthful, we kept them in the data when computing means and medians. [↑](#footnote-ref-11)
12. 41 people, in a previous question, said they were not born in the US, but then said in this question about documentation status that they were born in the US. So, the numbers here do not match Table 1, which gives generational status, because the variable “type” that we created as embedded data in Qualtrics was made using the “Were you born in the U.S.” yes/no question. Also, 32 respondents said they were born in US in the previous question, but then indicated some immigrant/related documentation status in this question. We need a coding rule to deal with those respondents whose answers to these two questions do not line up? [↑](#footnote-ref-12)
13. Is it odd that there are people currently in community colleges who already have associate’s degrees or higher. Not clear as to how best to code these inconsistent responses. [↑](#footnote-ref-13)