



Gutenberg School of Management and Economics
& Research Unit “Interdisciplinary Public Policy”

Discussion Paper Series

*Preference for Identification in the
Field - Nudging Refugees' Integration
Effort*

Nora Grote, Tim Klausmann, Mario Scharfbillig

March 11, 2019

Discussion paper number 1905

Contact details

Nora Grote
Johannes Gutenberg University
Gutenberg School of Management and Economics
Jakob-Welder-Weg 4
55128 Mainz
Germany

nora.grote@uni-mainz.de

Tim Klausmann
Johannes Gutenberg University
Gutenberg School of Management and Economics
Jakob-Welder-Weg 4
55128 Mainz
Germany

tim.klausmann@uni-mainz.de

Mario Scharfbillig
Johannes Gutenberg University
Gutenberg School of Management and Economics
Jakob-Welder-Weg 4
55128 Mainz
Germany

mario.scharfbillig@uni-mainz.de

PREFERENCE FOR IDENTIFICATION IN THE FIELD – NUDGING REFUGEES’ INTEGRATION EFFORT*

Nora Grote[†]
Tim Klausmann[†]
Mario Scharfbillig^{†‡}

March 11, 2019

Abstract

Social identity greatly affects behavior. However, less is known about individual’s preference for identification, i.e. how individuals choose their identity and more specifically whether and how subjects invest into belonging to a social group. We design a field experiment that allows us to make effort as an investment into a new group identity salient. The social identity in our treatment is refugee’s identification with the host society. We modified a mailing to 5600 refugees who use an online language-learning platform to learn the host countries’ language. These treatment emails make salient that improving the host country’s language ability increases the belonging to the host society. Our analysis reveals that the treatment has a significant positive effect on the effort exerted on the language-learning platform, leading to more completed exercises and more time spent learning the host country’s language. This suggests that refugees’ value being part of the host country’s society for its social identity component, which in turn reveals a general preference for identification. JEL codes: C93, D91, J15

*We would like to thank Steffen Altmann, Alexander Cappelen, Florian Hett, Yan Chen, Uri Gneezy, Daniel Schunk, Sigrid Suetens and Bertil Tungodden as well as the participants at the Spring School in Behavioral Economics in San Diego, the Workshop on Microeconomics in Lüneburg, the Current Topics in Public Sector Research Conference in Brno, the Microeconomics Seminar in Riezlern, the IMEBESS (2018) conference in Florence, the ESA World Meeting (2018), the 13th NCBE in Odense, the COMPIE (2018) in Berlin, the PhD course Intergenerational Economics in Frankfurt, the ZEW Workshop on the Economics of Immigration (2018) and the Advances with Field Experiments (2018) conference in Boston for helpful discussions and valuable feedback. We are grateful for the opportunity to work with the LinguaTV team and obtain their unique data. We received financial support from the research priority program Interdisciplinary Public Policy (IPP) at the University of Mainz.

[†]Johannes Gutenberg University of Mainz

[‡]Corresponding author: Jakob-Welder-Weg 4, 55128 Mainz, Germany, Phone: +49 6131 39 22918, Fax: +49 6131 39 27695, mario.scharfbillig@uni-mainz.de

I INTRODUCTION

Exploring further, the verb “to identify” is a necessary accompaniment of identity. There is something active about identity that cannot be ignored: it isn’t ‘just there’, it’s not a ‘thing’, it must always be established.

(Jenkins, 2014, p. 18)

Individuals’ social identity is an active process of self-identification. This process can take the form of an individuals’ investment into one or multiple social identities.¹ The willingness to invest depends on the social group(s) a person wants to belong to and wants to be considered to be part of by others (Jenkins, 2014). We call this need for belonging to a group the preference for identification. The membership in a social group has been shown to have an independent and substantial effect on behavior, such as public good contributions (Benjamin et al., 2016; Candelo et al., 2017; Charness et al., 2014), out-group discrimination (Hoff and Pandey, 2006), test performance (Benjamin et al., 2010; Hoff and Pandey, 2014), honesty (Cohn et al., 2014, 2010), altruism (Chen and Li, 2009) and cooperation (Chen et al., 2014). However, group identities employed in these studies are exogenous to the individual, such as ethnicity, gender or family. They may also be completely arbitrary or acquired later in life, such as membership in a platoon in the military (Goette et al., 2006). In contrast, in many contexts in life, identities are a matter of choice (Leary et al., 1986; Peirce, 1995; Stryker and Serpe, 1982; Tajfel and Turner, 1986). Even though the seminal paper on “economics and identity” (Akerlof and Kranton, 2000) pointed this out, the questions with which groups we identify and to which extent we invest into a process of identification, are largely neglected in the literature.

We seek to address the question whether individuals reveal a preference for identification by investigating whether and to what extent individuals invest into becoming part of a group in a migration context. Although, people may choose to which group they belong, identification may require some investment or change in behavior for it to be credible to the self and others. They may, for example, need to invest into group specific skills and

¹Identities are in nature multiple, i.e. it is possible to be at the same time white, female, young and a researcher. Studies investigating behavior specific to one of the multiple identities then make one dimension salient and interpret the behavior as stereotypical for that group (see Benjamin et al., 2010).

symbols, such as learning the language, reading (local) newspapers or wearing clothes in the fashion of the group. From observational data, it is often not possible to identify specific investments and behaviors as a sole means to become part of a group. In the case of language learning, the challenge is to separate the identity investment from multiple other reasons why immigrants may invest into learning the host country’s language. The most obvious would be the economic benefit through a higher probability of becoming employed and receiving higher wages (Chiswick and Miller, 1995; Dustmann et al., 2003; Lazear, 1999; Mcmanus et al., 1983).

We tackle this issue of disentangling the investments into identity from others by making use of a unique setting involving a large number of newly arrived refugees.² We designed a treatment, which only makes salient that investing into learning the host country’s language increases the sense of belonging to the host society, while keeping everything else equal. This increased salience will only lead to investment into language learning if the participants have a preference to identify with the host country’s society. Therefore, our treatment uses nudging in order to detect whether individuals exhibit a preference for identification.³ Furthermore, we can also observe the direction of our treatment. If refugees (do not) want to belong to the host society, then we would expect (less) more investment into learning the host country’s language.

By investigating the recent refugee wave, we seek to address the question whether individuals have a preference for identification with a policy relevant population. An important factor for the success of the integration process of refugees is learning the host country’s language, both as a means to learn about customs and as an end itself (Clots-Figueras and Masella, 2013; Hicks et al., 2015). Investigating changes in effort invested in learning a new language is particularly interesting, because it is a way to improve interaction efficiency. This allows for increasing acceptance within the host society, making

²By the term refugee we refer to a group of migrants that requested asylum in their recent country of residence independent of their current legal status. This group consists of people who are still awaiting a decision on their status as well as different granted legal statuses and rejected asylum applicants.

³Nudging, according to Hansen (2016) consists of “a function of any attempt to influence people’s judgment, choice, or behavior in a predictable way (1) that is possible because of cognitive boundaries, biases, routines and habits in individuals social decision-making posing barriers for people to perform rationally in their own declared self-interest and which (2) works by making use of those boundaries, biases, routines and habits as integral parts of such attempts” (Hansen, 2016, p.158)

it easier for the individual to identify and to be identified by others (Bauer et al., 2005; Dustmann, 1994; Georgiadis and Manning, 2013; Jenkins, 2014).

We design emails to be sent out to refugees already registered on an online language-learning platform. All emails informed them about the learning software being newly available as a mobile application as our neutral baseline. In the identity treatments, we additionally mentioned that learning the language helps to increase their belonging to the group. Adding this content on belonging makes sure that we manipulate the identity dimension only, enabling us to interpret our findings as being driven by a preference for identification. Moreover, this identity intervention comes at close to no cost as additional text in an email is not charged, which makes a cost-benefit calculation in case of a positive treatment effect easy.

More specifically, in one identity treatment branch the participants were informed in the subject line and body of the mailing that by learning German they could “become part of German community” (gain identity treatment), while in the other treatment branch it is stressed that they could “remain part of German community” (loss identity treatment). Apart from the paragraph in the email and the email subject, all participants, received the same email text. Therefore, the only differences between the groups is the salience of the fact that learning the host country’s language is an investment into their social group identity.

We find that our identity intervention succeeds in significantly increasing language-learning behavior. We find an increased likelihood of opening the email, clicking on the download and browser links, as well as more new logins, completed exercises, points achieved and more time spent on the platform after four weeks. Our results reveal relevant increases in learning, for example, learning time on average doubles relative to the control group after four and eight weeks. Our results suggest that refugees value being part of the host country’s society and that this induces them to exert more effort. Using post treatment survey data, we can show that our sample seems to be representative for the recent refugee wave in Germany for a large set of variables, with the exception that our sample is on average more educated. We do not find significant effects of our treatment on

labor market outcomes like employment probability. However, this seems to be reasonable given the limited sample size, the overall limited usage of the platform and the short time horizon (8 weeks). Finally, we also find no significant difference between the identity gain and identity loss treatment.

Our paper contributes to several strands of literature. First, we contribute to the empirical literature on identity in economics. By showing that individuals actively invest into their group identity, we provide evidence for endogenous identity formation and a preference for identification. This means that identities may be flexible and subject to individuals' history (Akerlof and Kranton, 2000; Jenkins, 2014; Tajfel, 1974, 1978; Tajfel and Turner, 1986). Our setting is especially relevant, because we are the first to experimentally investigate a "real" identity, which is not assigned by an experimenter.

A paper closely related to ours is Hett et al. (2017). They show that, using a modification of the minimal-group paradigm in the lab, individuals sacrifice monetary payoff in order to belong to a group based on group status. Status differences are induced by a quiz performance split, such that groups consisting of better performers may be more desirable. Heap and Zizzo (2009) show that people are willing to trade group memberships at a premium, from which they conclude that there exists a "psychological benefit" for groups in an interpersonal trust game setting. Both papers, however, are laboratory experiments using artificially assigned identities. Hence, it is questionable whether the results can be extrapolated to the real world. For example, in Hoff and Pandey (2006) the predictions from the identity literature on out-group punishment from the lab do not hold for lower caste members in India. Additionally, both Heap and Zizzo (2009) and Hett et al. (2017) measure identity choice using short-term monetary investments, while we look at investments in terms of effort over time. Our intervention further differs from prior work since we target an identity, which cannot be claimed by paying a fee. In lab experiments, individuals immediately switch groups in exchange for monetary payoff while in our setting the membership credibly depends on self-identification. This may very well cast doubt on the possibility to truly acquire an identity and therefore potentially reduce the willingness to invest into it, which we, however, show not to be the case. At the same

time, this setting is more realistic, as most identities are likely not purchasable.

Our research is also relevant to the growing literature strand on integration of immigrants and the role of identification with the host country’s society in particular. The effects of a migrants’ self-assigned identity on integration have received increasing attention in recent years (Constant et al., 2009; Epstein and Gang, 2010; Epstein and Heizler, 2015) and are found to be relevant to outcomes in the labor market (Bisin et al., 2016, 2011; Manning and Roy, 2009). However, this line of research mostly relies on survey evidence, correlating identification with important economic outcomes, such as labor market performance (Cameron et al., 2015). They usually find that higher identification with the host country’s society and societal values has a positive association. Some studies also find negative associations because subjects may choose oppositional identities, often summarized under the term of “acting white”, where actions towards others’ group norms are repellent (Austen-Smith and Fryer, 2005; Battu and Zenou, 2009; Fryer and Torelli, 2010). Our study provides causal evidence for a preference for identification. The findings suggest that on average refugees want to belong to the host country’s society, providing evidence against a widespread oppositional identity formation for this specific immigrant group in Germany.

Finally, we provide evidence that effort invested into a group identity can be nudged (Thaler and Sunstein, 2008). Despite the low cost nature of our identity intervention, with marginal cost of adding a few sentences to an email, we are able to significantly influence effort invested by refugees into learning the host country’s language. Therefore, our findings can inform policy for other integration measures aimed at immigrants in general.

II NUDGING LANGUAGE LEARNING EFFORT THROUGH PREFERENCE FOR IDENTIFICATION

In general, we use the term identity as a person's identification in a social group context, which originates from perceived membership in social groups, e.g., race, gender, nationality or religion (Tajfel, 1974; Tajfel and Turner, 1979, 1986). Literature on identity in economics is fast growing and was formally introduced by Akerlof and Kranton (2000). One strand is based on the minimal-group paradigm, sorting participants in the lab based on arguably unimportant categories such as preferences for different painting styles (e.g. Bernhard et al., 2016). Another strand makes genuine identities salient or primes them. Hoff and Pandey (2006), for example, show that cast affiliation, when publicly announced, decreases subsequent performance of low-caste members. Benjamin et al. (2010) prime either racial or gender identity using loaded terms in questionnaires and show that this has an influence on time and risk preferences. Differing from these types of studies, we make it salient to refugees that learning the host country's language is a way to change their sense of belonging to the host country's society.

An important question is, of course, why people would want to choose an identity, i.e. why they would have preferences for identification (Kranton et al., 2016). While we cannot sufficiently answer this question in our context, recent findings have proposed several channels for identity choice. Benabou and Tirole (2011) propose that people fundamentally value knowing "who they are" while at the same time not knowing the "correct" answer. Akerlof (2017) proposes that people have a valuation for self- and peer-esteem, which drives them to acquire certain identities and moral values. This esteem may be derived from the differential social status attached to a certain identity (Hett et al., 2017). In the context of migration, it seems likely that refugees know where they are from. Given their lost frame of reference in a new country and an often sudden break with the background, they might be inclined to acquire a new identity, which could potentially conflict with the origin identity (Constant et al., 2009). This new identification with the

host society may be perceived to having a higher social status within the new society, from which immigrants may obtain self-esteem.

Our sample drawn from the current refugee wave offers a good showcase on how preference for identification may be used to nudge effort in an immigration context in general. More than 1.3 million refugees arrived in Europe in 2015 from countries, such as Syria, Iraq and Afghanistan. Many of these refugees seek asylum in Germany (Brücker et al., 2016, 2017, 2018). A substantial fraction plans to stay for a longer time period. In fact, according to survey evidence, the majority of refugees seems to be highly motivated to become part of German society (Brücker et al., 2016). Nevertheless, for these refugees as for other migrant groups, it is difficult to assess, how this self-reported motivation translates into behavior.

Coercing immigrants into policies intended to increase their identification with the host country may also backfire due to the phenomenon of oppositional identities. Especially those immigrants with a poor socio-economic background, may well oppose trying to learn the language and behavior of the majority group. For example, African American students have been found to be ambivalent about learning “standard” English and to perform well at school because this may be interpreted as “acting white”, i.e. assimilating to the mainstream identity (Battu and Zenou, 2009; Bisin et al., 2011; Fordham and Ogbu, 1986; Fryer and Torelli, 2010; Ogbu, 1999; Patacchini and Zenou, 2016).

Overall, it is still unclear how exactly the process of identification with the host society works. Some immigrants may want to actively engage with the host country’s society, form relations and learn group-specific behavior such that they identify with it actively (Davis, 2007). Others, may just live in the country and consume identity-loaded products such as media and food and over time identify themselves with the host country’s society. Again others may simply retain their own identity because they value their own background, such as some Jewish communities in the US seem to do (Bisin and Verdier, 2000), or they may actively reject the majority identity as in the case of oppositional identities (Dehdari and Gehring, 2017; Monscheuer, 2018).

Our email treatment can be interpreted as a nudge towards a specific behavior that

refugees may or may not like and may react differently to (Thaler and Sunstein, 2008).⁴ We directly communicate with the users of a language-learning platform via email, prompting them to change their behavior depending on their preferences for identification. Importantly, this nudge can only work, if refugees have not previously optimized over their choice of identities and the investment into learning the language is not their best response already. In fact, in their study Hoff and Pandey (2014) conclude that the choice of identity is frame-dependent, which means that participants were not aware of all dimensions of their identity at all time. Therefore, making the identity-building component of language learning salient may change behavior.

Additionally, our identity intervention is framed as a loss for one third of our sample and as a gain for another third based on the vast literature showing the effect of reference dependent preferences, in which losses loom larger than gains (Kahneman Daniel and Tversky Amos, 2011; Tversky and Kahneman, 1991). There are numerous effects consistent with this finding, e.g. the endowment effect (Thaler, 1980). In field experiments, the additional motivation effect of loss framing is found to be successful in motivating teachers (Fryer et al., 2012; Levitt et al., 2016), marketing messages (Bertrand et al., 2010) or productivity (Hossain and List, 2012).⁵ We seek to investigate whether this finding also holds with regard to the identity dimension.

⁴In their review article, Damgaard and Nielsen (2017) conclude that in the field of education reminders frequently show robust positive effects. Additionally, reminders can be effective for real effort tasks and habit formation. Calzolari and Nardotto (2016) find that their reminders do not only lead to short term actions, but achieve real-effort outcomes. In a gym setting, they find signs of habit formation. In our context this finding would translate to continuous language learning instead of a short spike in learning only.

⁵While early results like Banks et al. (1995) find large advantages of loss-framing of health-related messages, recent work established small effects on health outcomes and overview articles show even opposing results. O’Keefe and Jensen (2009) argues in line with Kahneman Daniel and Tversky Amos (2011) and find small advantages of loss-framed messages compared to gain-framed ones. Gallagher and Updegraff (2012) find larger effects for gain-framed messages. Karlan et al. (2016) and Karlan et al. (2012) analyze savings behavior and cannot detect differences between loss- and gain-framing in reminder messages.

III STUDY DESIGN

III.A Participants and Context

During the height of the refugee influx in Germany in 2015, the language-learning platform we partnered with had donated several thousand licenses, worth 3-months of German learning, to the Federal Labor Employment Agency in Germany for distribution to refugees. The distribution of these licenses was neither random nor tracked in any way. In order to use the software, users had to register online with their email addresses. We accessed 5600 email addresses of refugees who received one of the donated licenses and registered on the learning platform. We composed emails, to be sent out to these refugees informing them about the platform being newly available as a mobile application. Before we sent our treatment emails, activity on the language-learning platform was low and 89% of the licenses were already expired.⁶ We extended or renewed all licenses for the time span of our identity intervention and observation period, which was two months.

On the platform, the language courses are organized similarly to a language textbook. Courses include visual, audio and textual elements. They consist of a large number of videos showing conversations in German. Each language lesson concludes with exercises and tests. The learner can repeat or skip lessons and follow her own speed of learning. The learning platform is available as a desktop version and at the time of our intervention newly as a mobile application, which differs from the desktop version only in the presentation style.

III.B Treatment

We designed three different emails sent to equally sized groups. We stratified over usage before the intervention and the date and time of registration to the platform. More specifically, we randomized triplet-wise, drawing three individuals and assigning them randomly to one of the treatments each in order to improve the balancing of our treatment assignment. Ex-post randomization checks confirm no significant differences in our

⁶Some users were able to renew their account by typing in the same donation key again, after it expired. Less than 1% of users did find this loophole and used it.

treatment and control groups (see table V in appendix).⁷

Our treatment consists of the email subject and content. One group (control group) received a baseline email which only informed participants about the new opportunity to learn on the mobile application and the two new months of free usage. A second group ("identity gain") received the same information but we added "Become part of the German community" at the end of the subject line and the following paragraph to the email:

Become part of the Germany community: Learning German will help you become part of the German society. It will allow you to connect with others and help you feel at home.

In a similar way, the third group ("identity loss") received an email including the following paragraph:

Stay part of the German community: Learning German will help you stay part of the German society. It will ensure you stay connected and do not feel isolated.

We sent these emails to the same individuals three times, whereby the second and third emails were marked as being reminders in the subject line. Additionally, for all three groups the second email contained "Learning German helps you find a job." and the third email contained "Learning German is important." The primary language of the emails was German, with translations to Arabic, Persian and English below. On top of the email, one could click on the name of each language to get directly to the corresponding part of the email. As almost all licenses were expired prior to our treatment and users only learned about the extension through our emails, we are confident that all effects work directly through our emails.

With the identity email treatments, we want to examine whether individuals increase effort towards attaining group membership. Therefore, we make salient that language learning may increase the identification with the host country's society. We chose our content such that we only mention aspects valuable within the context of belonging to that group, while we did not mention economic benefits, which may be valuable to individuals

⁷The difference for females is the only marginally significant variable. In our regressions we control for all available variables, including gender.

irrespective of their group membership. Economic benefits may nevertheless be implied, but the control group also receives an email reminding them of the language-learning opportunity, which itself implies a potential value to it. The only difference between the emails is the salience of a positive impact on group identification. Therefore, we believe that the salience of the economic dimension is not increased between control and treated groups (or between loss and gain). By directly connecting language learning and social identity, we are confident that the effect we capture is measuring the intended underlying concept.

It could further be questioned, whether our identity variation on a gain-loss dimension is valid. For refugees, who arrived 2-5 years ago, we pictured two possible scenarios: On the one hand these refugees arrived in the midst of the "Refugees Welcome"-Movement (Connolly, 2015). Hence, it seems likely that they experienced relatively intense initial interaction, for example, with volunteers, allowing them to feel well-integrated. From this perspective, improving language skills is central for keeping their gained social connections and thus their standing within the host society. On the other hand, in many situations they are still treated as refugees when receiving a donated license for the language-learning platform. Thus, learning behavior is an element of aspiring to gain a host culture identity. As discussed in O’Keefe and Jensen (2009) we keep the kernel state⁸ positive over both these methods to be able to detect the pure loss- versus gain-framing effect.

III.C Predictions

Given the survey evidence gathered from the recent refugee waves (Brücker et al., 2016), it seems plausible that most refugees in Germany do indeed want to be part of German society. Therefore, making it salient to them that learning the language increases their sense of belonging should nudge them into exerting more effort on learning German. However, the survey evidence could also be misleading because respondents do not have incentives to respond truthfully. On the contrary, knowing that this survey is prime

⁸“The kernel state is the basic, root state mentioned in the message’s description of the consequence.” (O’Keefe and Jensen, 2009, p.298) This simply means, that we refrain from using negations like "Not learning German will hinder your integration". We keep the action (learning) constant, while the consequence depends on the frame (become vs. stay part of Germany).

evidence on which policies may be designed later on, answering negatively may hurt refugees in the long-run. Nevertheless, we believe that coming to Germany is already evidence for a specific country choice. This leads to our primary hypotheses:⁹

H1: *An identity framed mailing increases investment into acquiring the identity of the host country.*

The treatment emails were either framed as a loss or a gain. In many domains losses loom larger than gains, which we expect to also be true for the identity dimension. Therefore, our second hypothesis is testing this:

H2: *Loss framed identity mailings have a stronger effect on investments into acquiring the identity of the host country compared to gain framed mailings.*

IV RESULTS

IV.A Data and Descriptive Results

Overall, we employ three sources of data. Firstly, we observe an individual’s activity on the language-learning platform. For a period of 3 month, we obtained cross-sectional data extracts from the backend of the platform. This data includes gender (self-assigned), registration dates on the platform, and information on the learning activity (number of logins, number of exercises started and points achieved per unit). Secondly, another data source stems from the mailings itself. We observe bounce rates, opening rates and click rates. The data on mailings and platform usage covers all 5600 individuals in our sample. Lastly, eight weeks after our intervention, we sent out an extensive questionnaire to all participants in our sample. We incentivized survey participation by a lottery offering prizes.¹⁰ This survey provided us with information on demographics, life events, employment history and status, ethnic identity and identification, locus of control and German skills for a subsample of our treatment population. The German skills were measured using a short and standardized German test developed by the Goethe Institute, the

⁹We preregistered these hypotheses on socialscienceregistry.org as Grote et al. (2017).

¹⁰The lottery was framed as a „gift handed out to some participants”, because of potential sensitivities of Muslims towards lotteries (Falk et al., 2017).

reference institution for German language learning. Table I shows descriptive statistics obtained from the backend of the language-learning platform and the mailings, results from our survey will be discussed in section IV.D First participants registered two years prior to our intervention. About one quarter (25.6%) of our sample is female.

[Table I about here]

In order to examine whether our treatments do affect learning behavior on the language platform, we will consider four different outcome measures. Firstly, we consider the number of logins to the platform indicating whether and how often the language learning platform was accessed. Prior to our intervention, participants on average logged in 20 times, whereby the median participant logged in six times and a few participants logged in more than 500 times. On the platform a learner can either solve exercises or study new elements by watching videos or listening to audio recordings. We consider the number of exercises an individual completed as well as the number of points an individual gained by solving exercises. The number of points received for working on an exercise depends also on the correctness of the answer and the complexity of the exercise. As a last indicator for learning behavior on the platform, we use learning time, which captures the amount of time an individual spent on the online language-learning platform in general. Learning time, logins and exercises are therefore measures for effort, while points additionally measure achievement. On average individuals completed 95 exercises, gained 5756 points and spent 6h and 45min on the platform before the start of our mailing intervention. 45% of the participants registered on the platform, but never started learning. In order to control for previous activity in our analysis, we create a dummy variable which is equal to one if the participants registered on the platform and completed at least one exercise (active before treatment). Activity during the four weeks prior to our intervention was very low. Only 194 individuals (3.5% of our sample) logged in at all and gained on average 141 points. This is consistent with previous findings that the use of voluntary online learning is low in general (Escueta et al., 2017), but is mainly due to the expired licenses in our sample.

IV.B Results of the Intervention

We estimate our effects using standard OLS regressions and heteroscedasticity-robust Huber-White standard errors I (MacKinnon and White, 1985). Additionally, we compute standard errors derived from randomization inference (Heß, 2017) that take into account the triplet-wise stratification we used. With these standard errors we do not rely on asymptotic properties of classic inference, making our analysis more robust. Our results are unchanged using this alternative error computation strategy. From the platform backend, we do have several observations over time for each measure. For brevity we report the effect of our intervention in regression analyses only for our final observation (8 weeks after the first email) and midway after four weeks. We therefore regress model (1), using only two points in time for identifying the treatment effect (McKenzie, 2012):

$$y_{i,post} = \alpha + \beta identity + \gamma y_{i,pre} + \sum_k \delta_k X_{i,k} + \epsilon_i \quad (1)$$

The variable *identity* is our treatment dummy variable. In order to investigate H1, we estimate our treatment effects first combined in one treatment dummy, and for H2 separately for the gain and loss treatment. Therefore, at first the dummy identity is equal to one, for an individual receiving any kind of identity-framed email. In order to control for the prior learning level, we also include the individual value of the dependent variable prior to our intervention where available. $X_{i,k}$ are k additional control variables, such as gender, month since registration on the platform and active before treatment.

Table II shows the immediate reactions to our identity intervention. We are able to directly observe whether an email was opened and whether a link in the email to download the application or open the platform in a browser was clicked.¹¹ Further, we observe whether and how often an individual logged into the program. Note that this is not necessarily the path participants need to go in order to learn more. Participants who keep the platform open in a browser window, for example, might just switch to it and others might react by visiting the platform later without using the link. Our emails were

¹¹This measure is not perfect, as we only observe the opening reaction of the email, depending on the email client and user settings. Therefore, this measure is a lower bound.

successfully sent to almost all of the registered users in our sample with a low bounce rate (2.55%). The first two columns of table II report the effect of the treatment email and its reminders combined compared to the control emails.¹² 66% of the participants opened at least one of our emails and 22.4% used the link in the mailing to reach the platform (see table I). This is an exceptionally high rate, compared to other studies using email interventions (e.g. Chen et al., 2017) and might indicate a very high interest in the language related emails.

We find that the identity treatment significantly increases the likelihood of an email being opened by 8.1 percentage points and the click rate by 2.3 percentage points. Clicking a link also translates into slightly more new logins at a marginally significant level. The identity treatment leads to 1.3 percentage points more logins. Given that overall 9.2% of our sample logged in within 8 weeks after our intervention, this implies 14% more logins in the identity treatments compared to the control treatment. The likelihood is also higher, if subjects were active before the treatment. Additionally, female subjects react less to our mailings in all measures. Overall, we conclude that the identity-framed subject line and email body induce more emails to be opened and also more clicks on the link as well as more logins into the platform.¹³

[Table II about here]

Next, we report our findings with regard to the main variables of interest, our measures for effort spent on language learning. Figure I shows the percentage of people having newly logged-in overtime and the mean per group for our learning behavior variables. In this graph, we differenced out prior activity.¹⁴ We observe that, the groups receiving the identity-framed email spent more time, earned more points and started more exercises

¹²When analyzing the first email and the two reminders separately, we see that only the first email had a significant treatment effect on both opening the email and clicking on the link, while the reminders were insignificant.

¹³We also tested for conditional effects of the treatment on clicking the link and logging in, if subjects opened the email. Again, our treatment could work, even if individuals do not open an email, due to the subject line. Hence, the decision to open the email may induced by the treatment itself and therefore could be endogenous. For clicking on a link in the email we find no significant treatment effects but we do find that the probability of logging in at least once during 4 weeks after the first email increases by 2.4% in the treatment conditional on having opened the email (see table A.3).

¹⁴We subtract the value of a variable at the time of the first email sent for all measures which have prior values, such as learn time, exercises and points earned.”

than the control group.¹⁵ The extent to which participants in loss and gain identity treatment groups react to our mailings is very similar albeit individuals in the loss treatment group seem to react slightly stronger.

[*Figure I about here*]

For a more precise estimates of treatment effects, we estimate the effects in a regression using model (1) controlling for time since registration, activity on the platform before our intervention, gender, and the pre-intervention level of the dependent variable. The first three columns in table III report the effect on learning behavior after 4 weeks, the last three columns report the results measured 8 weeks after our first email was sent. We consistently find positive effects on all measures available. After 4 weeks, the treatment effect with respect to learnTime increases by 5.4 minutes on average (significant on the 1% level), the increase considering number of exercise and total points as outcomes is marginally significant. After 8 weeks, the only dependent variable where we can still measure a significant increase caused by the identity intervention is learnTime (8.3 minutes more time learned on the platform), while the treatment effect on other measures is positive with larger standard errors compared to the observations after 4 weeks. Taken together, we see that identity framed emails do have a positive significant effect on effort spent learning the host country’s language. However, for the the measure more closely resembling achievement (points) the effect is also positive but seems to fade out over time.¹⁶

[*Table III about here*]

Regarding H1, the absolute effect size of our treatment is modest. Nevertheless, given that overall usage of the platform is very low prior to our treatment, the increase in activity through the identity framed emails is meaningful. During the first 4 weeks after our identity intervention, treated individuals learn on average 2.1 times as long as individuals

¹⁵The recording of the variable learnTime was changed two weeks prior to our intervention by our cooperation partner, which prevents us from showing a prior trend of it.

¹⁶The marginal effects of our treatment at all other points in time are summarized in Figure A.1 in the appendix. We also report robustness checks using winsorizing as well as trimming (both 1st and 99th percentile) in Tables A.7 and A.8.

in the control group (9.9 min vs. 4.6 min) and complete 1.7 times as much exercises (3.9 vs 2.3, see table A.2 in the appendix).

A potential limitation of our finding may be that making language learning salient as an investment into identity could subtly also convey information on other dimensions unrelated to identity. Refugees arriving in a new country arguably receive a lot of new information, due to the new surroundings, norms and customs. It is likely that they are not aware of which channel may be the best way of integrating, even if it is only about economic integration. By sending them a message with information on language learning, we make salient that learning German is one way to do so. This may nudge them to learn more German, which may not directly be related to identity. In our control treatment, however, we included all information except for the identity dimension. We therefore think this is unlikely for another channel to drive the results, but cannot exclude alternative explanations completely. Further, individuals who are only interested in economic integration and not identification with the host society, are also more likely to think instrumentally about coming to Germany. They are also most likely to have previously optimized their migration decision, i.e. they have chosen to leave their country behind and have chosen Germany as their best option as host country, given their abilities and preferences. Then, it seems unlikely that they are easily influenced to learn more German for economic benefits by the added section and the altered subject line in the email sent to them.

We now turn to the analysis of the differences between our gain and loss identity framing. In the analysis, we now include gain and loss identity treatment dummies separately. Columns 1-3 in table IV report the results for our mailings. The coefficients remain virtually the same for all measures regarding the direct response to the mailing, although in the gain treatment the coefficients for clicked (on link in email) and new login into the platform are not significant anymore. Columns 4-6 report the findings for our effort measures 4 weeks after the intervention. The coefficients of the loss treatment are numerically larger than the gain treatment and for the gain treatment all coefficients except for learning time are insignificant, while for the loss treatment we find all of them

to be significant (table A.4 in the appendix also reports the results 8 weeks after our intervention). Nevertheless, when testing for difference between the coefficients, we again cannot reject the null that both coefficients are equal for each measure. We conclude that for identity-framed mailings in a migration setting, gain and loss framing does not seem to matter much.

Overall, we do not find sufficient support for H2. There are several potential reasons for this. The way we framed the loss in our email may not be sufficient to trigger loss aversion. Alternatively, our sample of newly arrived refugees is too selected to react to it, in that they may not sufficiently identify with the German society, and therefore do not react to the potential loss. The findings may also result from the language barrier in the emails itself. There is evidence that loss framing in a foreign language has no different effect from gain framing (Keysar et al., 2012). Even though we translated the emails, the subject line was only in German and the first paragraph was also written in German. In order to read the email in Arabic or Farsi, individuals had to scroll down to their language or click on the respective in-text link.

[Table IV about here]

IV.C Channels

In this section, we focus on potential drivers of our main treatment effect for the preference for identification. We investigate whether prior activity and the language, in which refugees most likely read our emails in, influence learning behavior.

First, we interact the treatment dummy with prior usage of the platform.¹⁷ If a subject values being part of the German society, the motivation to use the platform should be higher on average and they should also react stronger to an identity treatment. As reported above, only 55% of registered users used the platform actively for learning before. Because we randomly distributed subjects into treatments, we can interact the information on prior activity with the treatment. If the effect of this interaction of prior

¹⁷For our interaction estimations, standard errors from randomization inference are not available.

usage and our treatment is significantly positive, it seems likely that we capture a valuation of refugees for being part of German society.

Tables V reports our findings for email and learning activity. As reported before, both the identity mailing and prior activity positively affect the likelihood of opening the email, clicking on the link and logging into the platform. We do not find any significant interaction effect on mailing reactions. Turning to our effort measures, we observe that the interaction effects are positive and significant for all measures (table A5 in the appendix also reports the results 8 weeks after our intervention). In contrast, the effect of the treatment dummy itself is not significant anymore and substantially closer to zero. For participants active on the platform prior to our intervention, either the treatment within the mail was the important part of the treatment inducing them to learn more, or they were sufficiently reminded to go to the learning platform by seeing the subject line itself, without needing to open the mail (significantly more). Overall, it seems that our treatment induced all individuals to open the mail and click on the link equally, while only the prior active individuals, who have arguably the highest incentive to invest into identity, also learned more in response to our treatment. We interpret this as evidence that our treatment succeeded in inducing participants to consider investing more into their identity, rather than a mere automatic effect on the platforms activity through a higher mail opening rate.

[Table V about here]

One specific concern in regards to identity in the literature is that of opposing identity. Some African Americans in the us seem to avoid learning more, because they do not want to seem as “acting white” (Austen-Smith and Fryer, 2005; Fryer and Torelli, 2010). We interpret our positive treatment effect together with the evidence from above from the interaction regressions as evidence against opposing identities within our sample, at least on average. However, due to a lack of a big enough group of individuals still active at the time of our intervention, we cannot investigate whether some individuals oppose the identity treatment by reducing their use of the platform. Therefore, we cannot completely exclude that subgroups of our sample behave in an opposing way regarding our treatment.

Next, we try to further decompose which individuals reacted to our treatment. In our emails, we tracked within which language block of the email the corresponding link to the language-learning platform or the app was clicked. Hence, we can proxy which language individuals were reading the email in. Similar to the argument before, individuals with a greater preference for identifying with the host country’s society would be more prone to click on the German link, either because they are already better in German when receiving our email, or because they are more willing to read something in German. In table VI, we interact the link language with our treatment effect. The important line is the interaction of our treatment with Clicked and German link clicked. Except for new logins, we find strong and significantly positive effects for the interaction, corroborating the idea that the result is driven by those individuals with the highest probability to care about the host countries’ identity.¹⁸

[Table VI about here]

Overall, the interaction effects with prior activity and the analysis of the clicking language indicate that the effects we find, are in line with preferences for identification. Compared to a neutral email, receivers of an identity treatment email learn more, especially when having learned more prior to our intervention and if they clicked on the German link showing a predisposition to learn and read German.

We further test if the preference for identification is gender specific. Table A.6 in the appendix reports our results where the treatment effect is interacted with gender. A significantly negative coefficient of this interaction effect would be exploratory evidence that females do either have a weaker preference for identification, or at least are more difficult to be nudged in this context. We find a significant negative interaction only for mail opening but of considerable size. Treated females are 7% less likely to open our email. The effects on all other variables are insignificant, but for the investment measures interestingly even positive. This means that even though they opened the email less,

¹⁸The tracking links in our emails also permit us to test, if the reading language is indeed limiting the effect of loss aversion (Keysar et al., 2012). We repeat analysis from table VI, interacting our treatment and the language of the link an individual clicked on. We assume that every language clicked on that is not German is the mother tongue of that person. In unreported results, we find no significant difference between the two different interaction effects on different languages for any of our measures used.

learning did not seem to be affected detrimentally. In unreported tests, we also estimate the interaction effect conditional on having opened the mail, to see if for that subsample the effect would be significant. Again, remember that the effect of our mails on learning could come directly through the subject line without the need for opening the email. We do not find any significant interaction. Overall, because our variables measuring investment are of prime focus in this investigation, we conclude that the preference for identification we detected might be a universal feature regarding gender.

IV.D Results from the Survey

Next, we focus on the data we collected from the survey that was sent to all participants 8 weeks after our intervention. 8.9% ($n=496$) of our original sample did fill out the survey completely (14% started the survey), which is quite a high number for an online survey, but nevertheless sharply reduces our sample size. The small sample size together with few users of the learning platform unfortunately prevents us from further looking into the structure of our treatment effect with other interactions. Nevertheless, the survey allows us to present some evidence that our subject pool is indeed comprised of recently arrived refugees, check for the representativeness and report potential effects of our treatment on labor market outcomes. For balancing, we checked whether participants in the identity treatments are more (or less) likely to complete our survey but could not detect any differences (coef -0.0052, p-value 0.526), which indicates that survey response is independent of treatment status.

First, we report socio-demographics for all non-missing responses. Although survey respondents do not necessarily need to be representative for the complete sample, they give us a general impression of the population at question. An overview of the participants' characteristics is shown in table VII. We coded no answers as missing. The mean age of our sample is 30, with a range from 12 to 63 years. Most people in our survey come from Syria (71.4%), followed by Afghanistan (5%), Iran (4.8%) and Iraq (3.3%). Around 89.3% already applied for asylum within Germany, and for 77% of them the requests have already been decided upon. A majority has an accepted refugee status (48%), is recognized as

eligible asylum seeker (30.2%) or received subsidiary protection (16.2%). Only a small amount of respondents has either a status of rejection without deportation (3.4%) or with deportation (2.4%). The vast majority of all respondents does eventually want to acquire German citizenship if they are allowed to (98.9%). In terms of education, our sample seems to be rather well educated, with 91.4% having completed at least secondary.

[Table VII about here]

We also asked participants to self-assess their German language ability in reading, writing and speaking on a five point likert-scale, and 41.6% of respondents answered naming one of the two highest options (“well” or “very well”), with minor differences between reading, writing and speaking. Additionally, we administered a German test at the end of the survey. This test was taken from a standard language level categorization test from the Goethe Institute, which is Germany’s cultural institute promoting German language learning worldwide. The test consisted of 30 questions separately testing reading, listening and writing skills. On average respondents solved 21.9% of the questions correctly which corresponds to the lower limit for the A2 level of the Common European Framework of Reference for Languages, described as “Can understand sentences and frequently used expressions related to areas of most immediate relevance”. Even the best respondent answered only 60% of the questions correctly, which corresponds to a B2 level (“Can understand the main ideas of complex text on both concrete and abstract topics”)¹⁹. This means that at least measured by this standardized short test, German language abilities of all survey respondents were low and did not perfectly correspond to the self-assessment.²⁰ Even though, the incongruence between the self-reported abilities and the test may partially be explained by the fact that the language test was at the end of our survey and that it was not incentivized, we take this as overwhelming evidence that almost all respondents could benefit from learning German with the help of the online-learning program.

¹⁹For more information on language qualification assessment grid, see https://en.wikipedia.org/wiki/Common_European_Framework_of_Reference_for_Languages.

²⁰We also checked, whether respondents simply click through the test and did not find evidence for this. Excluding the answer times of over 3 hours (which might be due to bad browser timeouts), the average respondent completed the language test part of the survey in 35 minutes (whole survey in 64 min).

[Table VIII about here]

In terms of current occupation, only few responded to be “waiting” (14.3%). Many respondents report to either go to integration courses (36.4%), to school/university (27%), to work (20%) or to complete an apprenticeship (8.1%) or internship (7.2%). This implies that for more than 85% of our sample, time to learn the language on an online platform may be limited or only complementary to traditional ways to learn the language.

We compared our survey answers with those from the representative IAB-BAMF-SOEP refugee panel for the recent refugee wave in Germany (Brücker et al., 2018), and find that our sample is in almost all respects very similar and hence, representative for the recent refugee wave in Germany. In terms of age, country composition, arrival time, and motivation to come, our sample closely matches the representative IAB-BAMF-SOEP refugee sample. We only observe strong differences in terms of education, with our sample being more highly educated. In our sample, around 90% claim to have finished at least secondary school, while in the IAB-BAMF-SOEP Panel, only 35% finished secondary school. In general, our sample is even better educated than the average in Syria. According to Morrisson and Murtin (2009), the average years of education in Syria is 8 years for the whole population, while in our sample the average is more than 13 years.

In table VIII we report the results of regressions of our treatment dummy on labor market outcomes. Influence in this dimension was not the main target of our investigation and given the overall low level of usage of the platform it is unlikely to have a strong influence. Nevertheless, our emails might have induced individuals to be more active in, for example, sending out CV’s. Column (1) reports the results for working, a dummy variable equal to one, if a subject responds to having a job. Column (2) reports net wages, column (3) weekly working hours (only for individuals working at the moment) and columns (4) and (5) report the number of applications and interviews an unemployed learners made in the last months. All effects are insignificant and close to zero, which does not seem to be surprising given the short time elapsed since the treatment, the low number of observations and the overall small effect of the treatment. In untabulated results, the effect remains insignificant, viewed separately in gain and loss or interacted

with prior activity.

V CONCLUSION

We investigate the extent to which people exhibit a preference for identification, measured by refugees' willingness to invest into the social identity of their host country. By using an experimental setting, we exogenously vary the salience of the investments' link to identification. Our results suggest that individuals do actively seek to invest into group specific identity. To the best of our knowledge, we therefore are the first to provide causal evidence from the field that supports the claim that people do indeed have a preference for identification.

We make use of a sample of recently arrived refugees in Germany. With this we are able to show that refugees can be motivated to invest more into learning the host country's language by using a nudge in emails, which make salient that learning the language is an investment into their identity. Compared to receivers of the neutrally-framed email, receivers of an email making identity salient complete more exercises and spend more time on the language-learning platform. The effects are stronger for refugees with more prior activity and for those, who read the email in the host country's language, making it plausible that our treatments measure a preference for identification. We do not find support for a stronger effect of loss framing relative to the gain framing identity, which may be due to a foreign language effect.

On average we do not observe evidence for oppositional identities in our sample. By analyzing the average response to the treatment we find that refugees in Germany respond positively to identity framed emails, increasing their learning effort significantly, which leads us to reject oppositional identities in our setting. However, we cannot look into individual responses, which might look different.

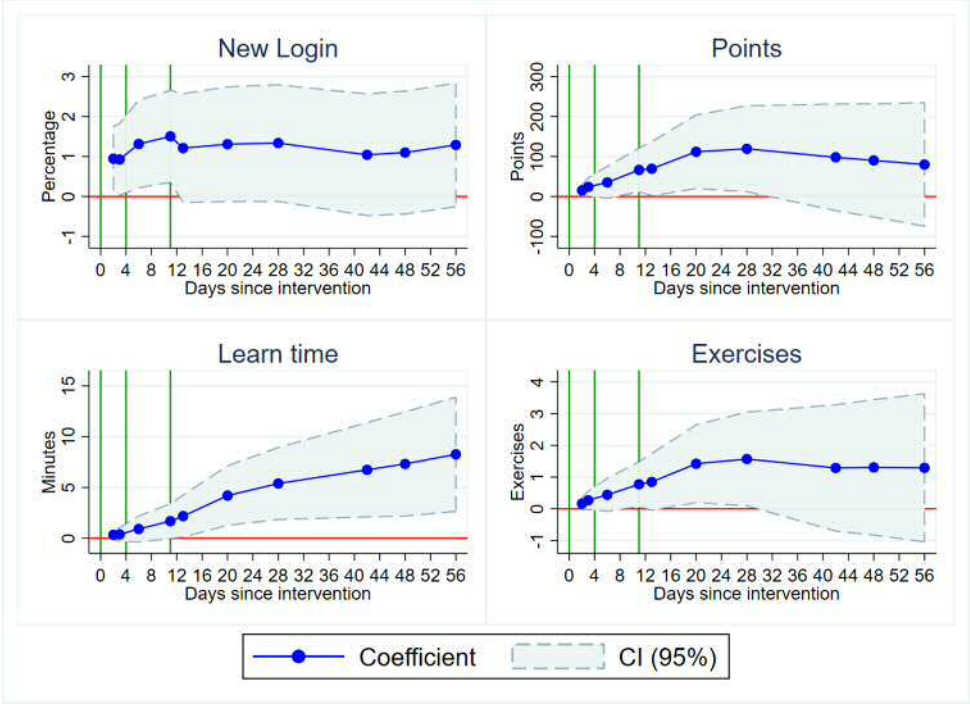
Our results imply that it may pay off to communicate using identity loaded messages with newly arrived immigrants in order to increase their effort spent on integration. This is in line with other research using identity-building activities, such as attending a citizenship ceremony (Manning and Roy, 2009). Our intervention comes at almost no costs as it only

requires the addition of a few sentences to communication. This makes it a very low-cost potential policy tool for increasing integration effort.

APPENDIX A - ADDITIONAL ANALYSIS

Figures

Figure A.1: Coefficients of Learning Behavior Over Time



Tables

Table A.1: Identity Treatment and Control Group Means and Difference Test

	(1) Treatment	(2) Control	(3) Δ and t-test
New logins	20.31 (40.03)	20.69 (49.40)	0.38 [0.758]
Exercises	95.64 (266.83)	93.59 (274.17)	-2.05 [0.789]
Points	5809.17 (16091.16)	5648.56 (16411.96)	-160.61 [0.727]
Learn time	405.19 (1099.19)	405.84 (1224.76)	0.65 [0.984]
Days since registration	401.91 (132.48)	401.77 (132.55)	-0.14 [0.971]
Active before treatment (d)	0.55 (0.50)	0.55 (0.50)	-0.00 [0.992]
Female (d)	0.26 (0.44)	0.24 (0.43)	-0.02 [0.051]*
Observations	3715	1858	5573

Note: Column (1) and (2) report the means of regression-relevant variables for treatment and control group recorded before the intervention and their standard deviations in parenthesis. Column (3) reports difference of means and p -values for two-sided t-test for mean difference in boxy parenthesis.

* $p < 0.1$, ** $p < 0.05$, *** $p < 0.01$. Standard errors in parentheses.

Table A.2: Regression Relevant Descriptives by Treatment

	N	All			Treatment			Control		
		Mean	SD	Min	Max	N	Mean	SD	Min	Max
Months since registration	5573	13.176	4.344	2	24	3715	13.177	4.344	2	24
Active before treatment (d)	5573	0.549	0.498	0	1	3715	0.549	0.498	0	1
Female (d)	5572	0.256	0.436	0	1	3715	0.264	0.441	0	1
<i>Mailings</i>										
Opened (d)	5570	0.660	0.474	0	1	3713	0.687	0.464	0	1
Clicked (d)	5434	0.224	0.417	0	1	3619	0.232	0.422	0	1
<i>Learning behavior before treatment</i>										
Exercises	5573	94.953	269.275	0	3481	3715	95.636	266.826	0	3481
Points	5573	5755.625	16197.524	0	215057	3715	5809.171	16091.164	0	215057
Learn time	5573	405.410	1142.477	0	15517	3715	405.193	1099.189	0	15276
<i>Learning behavior 4 weeks after treatment</i>										
New login 4 weeks (d)	5573	0.081	0.273	0	1	3715	0.086	0.280	0	1
New Exercises after 4 weeks	5573	3.332	31.269	0	1018	3715	3.864	35.046	0	1018
New Points after 4 weeks	5573	243.022	2348.949	0	74201	3715	283.453	2681.632	0	74201
New learn time after 4 weeks	5573	8.200	80.886	0	2203	3715	9.992	94.716	0	2203
<i>Learning behavior 8 weeks after treatment</i>										
New login 8 weeks (d)	5573	0.092	0.289	0	1	3715	0.096	0.294	0	1
New Exercises after 8 weeks	5573	4.952	43.826	0	1136	3715	5.393	45.005	0	1083
New Points after 8 weeks	5573	342.129	3001.620	0	77002	3715	369.619	3174.700	0	77002
New learn time after 8 weeks	5573	13.247	128.121	0	4042	3715	15.992	149.892	0	4042

Table A.3: Reactions to Identity Treatment Conditional on Opened Email

	(1)	(2)	(3)	(4)	(5)
VARIABLES	Clicked (d)	New login 4 weeks (d)	Exercises	After 4 weeks Points	Learn time
Identity treatment (d)	-0.000 (0.017)	0.024** (0.010)	8.374 (9.697)	572.565 (578.807)	12.547 (43.914)
Months since registration	0.000 (0.002)	-0.009*** (0.001)	-1.677 (1.121)	-96.966 (67.955)	14.013*** (4.160)
Active before treatment (d)	0.030* (0.016)	0.096*** (0.009)	183.684*** (7.949)	11,188.585*** (479.763)	796.593*** (34.465)
Female (d)	-0.065*** (0.017)	-0.021* (0.011)	48.062*** (12.454)	2,744.445*** (736.883)	109.129** (49.898)
Constant	0.314*** (0.029)	0.159*** (0.020)	6.717 (15.623)	373.221 (944.997)	-217.671*** (63.379)
Observations	3,629	3,677	3,677	3,677	3,677
R-squared	0.004	0.042	0.107	0.108	0.105

Note: Coefficient estimates from ordinary least squares estimations. * significant at 10%; ** sign. at 5%; *** sign. at 1%. The sample used in this analysis only contains learners who opened at least one of the three emails. The dependent variable is a dummy variable that takes the value of 1 when the learner clicked at least once on a link to the language- learning platform presented in the intervention emails, when the learner logged onto the learning platform at least once within four weeks of the intervention in column 2 and when the learner logged onto the learning platform at least once within eight weeks of the intervention in column (3). The explanatory variable of main interest is identity treatment which takes the value of 1 if learner i participated in the identity treatment and is 0 otherwise (control group - reminder only). Days since registration measures the time a learner was registered on the platform before the first intervention email was sent. Active before intervention is a dummy variable that takes the value of 1 when a learner spent time on the platform before the intervention and 0 otherwise. The dummy variable Female takes the value of 1 when learner i is female and is 0 otherwise. Heteroscedasticity-robust Huber-White standard errors are in parentheses.

Table A.4: Effect of Gain and Loss Treatment - Mailing and Learning Outcomes Full

VARIABLES	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)	(10)
	Opened (d)	Clicked (d)	New login (d)	Exercises	After 4 weeks Points	Learn time	New login (d)	Exercises	After 8 weeks Points	Learn time
Identity gain treatment (d)	0.082*** (0.016)	0.023 (0.014)	0.011 (0.009)	1.128 (0.852)	108.285 (72.345)	4.155** (1.917)	0.010 (0.009)	0.646 (1.305)	41.820 (92.058)	6.584* (3.427)
Identity loss treatment (d)	0.081*** (0.016)	0.023* (0.014)	0.016* (0.009)	2.013** (1.015)	130.282* (67.330)	6.626** (2.726)	0.016* (0.009)	1.942 (1.472)	118.014 (95.846)	9.945** (4.050)
Months since registration	-0.001 (0.001)	0.000 (0.001)	-0.009*** (0.001)	-0.762*** (0.164)	-48.310*** (10.665)	-2.466*** (0.491)	-0.010*** (0.001)	-0.966*** (0.208)	-60.301*** (13.569)	-3.368*** (0.664)
Active before treatment (d)	0.038*** (0.013)	0.025** (0.011)	0.079*** (0.007)	1.986*** (0.693)	145.590** (60.028)	5.665*** (1.419)	0.086*** (0.007)	3.217*** (0.978)	212.353*** (73.917)	8.102*** (2.700)
Female (d)	-0.048*** (0.015)	-0.060*** (0.012)	-0.020** (0.008)	-0.316 (1.002)	-22.035 (77.452)	-0.575 (2.510)	-0.023*** (0.008)	-0.872 (1.320)	-45.322 (95.964)	-1.032 (4.012)
Dep. var. before interv.				1.015*** (0.003)	1.015*** (0.004)	1.010*** (0.002)		1.022*** (0.005)	1.021*** (0.005)	1.018*** (0.005)
Constant	0.608*** (0.024)	0.210*** (0.021)	0.147*** (0.015)	9.939*** (2.117)	641.064*** (139.818)	29.970*** (6.169)	0.170*** (0.016)	13.239*** (2.705)	856.883*** (175.703)	40.538*** (8.192)
Observations	5,569	5,433	5,572	5,572	5,572	5,572	5,572	5,572	5,572	5,572
R-squared	0.010	0.005	0.041	0.988	0.980	0.995	0.045	0.976	0.969	0.989

Note: Coefficient estimates from ordinary least squares estimations. * significant at 10%; ** sign. at 5%; *** sign. at 1%. The dependent variable is a dummy variable that takes the value of 1 when the learner opened at least one of the intervention emails in column (1), when the learner clicked at least once on a link to the language-learning platform presented in the intervention emails in column (2). The dependent variables report what learners achieved on the learning platform in the first four weeks after the first intervention email was sent in columns (3) to (6) and in the first eight weeks after the first intervention email was sent in columns (7) to (10). The dependent variables report whether the learner logged onto the learning platform at least once in columns (3) and (6), how many exercises a learner completed in columns (4) and (8), how many points the learner collected in columns (5) and (9) and how much time the learner spent in columns (6) and (10). The explanatory variables of main interest are the two versions of the identity treatment. Both the identity gain and identity loss treatment variables take the value of 1 if learner i participated in that specific treatment and 0 otherwise (control group - reminder only). Days since registration measures the time a learner was registered on the platform before the first intervention email was sent. Active before intervention is a dummy variable that takes the value of 1 when a learner spent time on the platform before the intervention and 0 otherwise. The dummy variable Female takes the value of 1 when learner i is female and is 0 otherwise. Heteroscedasticity-robust Huber-White standard errors are in parentheses

Table A.5: Interaction with Prior Activity Learning

VARIABLES	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)	(10)
	Opened (d)	Clicked (d)	New login (d)	Exercises	Points	Learn time	New login (d)	Exercises	Points	Learn time
					After 4 weeks				After 8 weeks	
Identity treatment (d)	0.095*** (0.021)	0.028* (0.017)	0.004 (0.008)	0.255 (0.379)	23.306 (35.363)	-0.183 (0.718)	0.005 (0.009)	-0.171 (0.579)	-18.546 (53.947)	-0.259 (0.918)
Identity treatment * Active before treatment (d)	-0.025 (0.027)	-0.010 (0.024)	0.017 (0.014)	2.401* (1.388)	175.135* (102.000)	10.171*** (3.326)	0.015 (0.015)	2.673 (2.195)	179.658 (146.958)	15.554*** (5.253)
Months since registration	-0.001 (0.001)	0.000 (0.001)	-0.009*** (0.001)	-0.762*** (0.164)	-48.320*** (10.662)	-2.467*** (0.491)	-0.010*** (0.001)	-0.966*** (0.208)	-60.307*** (13.570)	-3.369*** (0.663)
Active before treatment (d)	0.055** (0.023)	0.031* (0.019)	0.068*** (0.011)	0.383 (0.933)	28.920 (60.245)	-1.121 (2.280)	0.076*** (0.012)	1.431 (1.599)	92.479 (102.418)	-2.274 (4.323)
Female (d)	-0.047*** (0.015)	-0.060*** (0.012)	-0.021*** (0.008)	-0.350 (1.000)	-24.361 (77.272)	-0.713 (2.520)	-0.024*** (0.008)	-0.910 (1.314)	-47.843 (95.503)	-1.243 (4.019)
Dep. var. before interv.				1.015*** (0.003)	1.015*** (0.004)	1.010*** (0.002)		1.022*** (0.005)	1.021*** (0.005)	1.018*** (0.005)
Constant	0.599*** (0.026)	0.207*** (0.022)	0.153*** (0.015)	10.825*** (2.262)	705.700*** (148.609)	33.726*** (6.777)	0.176*** (0.016)	14.225*** (2.861)	923.164*** (189.258)	46.282*** (9.094)
Observations	5,569	5,433	5,572	5,572	5,572	5,572	5,572	5,572	5,572	5,572
R-squared	0.010	0.005	0.041	0.988	0.981	0.995	0.045	0.976	0.969	0.989

Note: Coefficient estimates from ordinary least squares estimations. * significant at 10%; ** sign. at 5%; *** sign. at 1%. The dependent variable is a dummy variable that takes the value of 1 when the learner opened at least one of the intervention emails in column (1), when the learner clicked at least once on a link to the language-learning platform presented in the intervention emails in column (2). The dependent variables report what learners achieved on the learning platform in the first four weeks after the first intervention email was sent in columns (3) to (6) and in the first eight weeks after the first intervention email was sent in columns (7) to (10). The dependent variables report whether the learner logged onto the learning platform at least once in columns (3) and (6), how many exercises a learner completed in columns (4) and (8), how many points the learner collected in columns (5) and (9) and how much time the learner spent in columns (6) and (10). The explanatory variable of main interest is the identity treatment interacted with Activity before treatment which takes the value of 1 if learner *i* participated in the identity treatment and was active before our intervention and is 0 otherwise. Days since registration measures the time a learner was registered on the platform before the first intervention email was sent. Active before intervention is a dummy variable that takes the value of 1 when a learner spent time on the platform before the intervention and 0 otherwise. The dummy variable Female takes the value of 1 when learner *i* is female and is 0 otherwise. Heteroscedasticity-robust Huber-White standard errors are in parentheses.

Table A.6: Regression Interaction with Female

VARIABLES	(1)	(2)	(3)	(4)	(5)	(6)
	Opened (d)	Clicked (d)	New login (d)	Exercises	After 4 weeks Points	Learn time
Identity treatment (d)	0.098*** (0.016)	0.026* (0.014)	0.015* (0.009)	0.899 (0.891)	73.936 (64.536)	4.619** (2.152)
Identity treatment * Female (d)	-0.069** (0.032)	-0.013 (0.026)	-0.006 (0.017)	2.714 (1.669)	183.286 (125.155)	3.119 (4.004)
Months since registration	-0.001 (0.001)	0.000 (0.001)	-0.009*** (0.001)	-0.761*** (0.164)	-48.213*** (10.660)	-2.465*** (0.492)
Active before treatment (d)	0.039*** (0.013)	0.025** (0.011)	0.079*** (0.007)	1.948*** (0.692)	143.193** (60.004)	5.625*** (1.418)
Female (d)	-0.001 (0.027)	-0.052** (0.021)	-0.016 (0.013)	-2.165** (0.969)	-146.825** (66.717)	-2.701 (2.271)
Dep. var. before interv.				1.015*** (0.003)	1.015*** (0.004)	1.010*** (0.002)
Constant	0.597*** (0.024)	0.208*** (0.022)	0.146*** (0.015)	10.376*** (2.145)	670.615*** (140.668)	30.475*** (6.089)
Observations	5,569	5,433	5,572	5,572	5,572	5,572
R-squared	0.011	0.005	0.041	0.988	0.981	0.995

Note: Coefficient estimates from ordinary least squares estimations. * significant at 10%; ** sign. at 5%; *** sign. at 1%. The dependent variable is a dummy variable that takes the value of 1 when the learner opened at least one of the intervention emails in column (1), when the learner clicked at least once on a link to the language-learning platform presented in the intervention emails in column (2) and when the learner logged onto the learning platform at least once within four weeks of the intervention in column (3). The dependent variables report what learners achieved on the learning platform in the first four weeks after the first intervention email was sent in columns (4) - (6). The explanatory variable of main interest is the identity treatment interacted with Activity before treatment which takes the value of 1 if learner i participated in the identity treatment and was active before our intervention and is 0 otherwise. Days since registration measures the time a learner was registered on the platform before the first intervention email was sent. Active before intervention is a dummy variable that takes the value of 1 when a learner spent time on the platform before the intervention and 0 otherwise. The dummy variable Female takes the value of 1 when learner i is female and is 0 otherwise. Heteroscedasticity-robust Huber-White standard errors are in parentheses.

Table A.7: Learning Behavior - Winsorizing (1th and 99th Percentile)

VARIABLES	(1)	(2)	(3)	(4)	(5)	(6)
	Exercises	After 4 weeks Points	Learn time	Exercises	After 8 weeks Points	Learn time
Identity treatment (d)	1.191 (0.858)	112.847* (63.730)	4.354** (2.080)	1.401 (1.142)	98.748 (79.480)	6.060* (3.094)
Months since registration	-0.688*** (0.093)	-42.191*** (6.940)	-2.421*** (0.227)	-0.908*** (0.124)	-55.826*** (8.656)	-3.299*** (0.338)
Active before treatment (d)	-0.277 (0.869)	96.786 (64.557)	1.689 (2.121)	0.803 (1.157)	99.299 (80.512)	2.125 (3.154)
Female (d)	-0.308 (0.932)	-24.430 (69.235)	-0.250 (2.258)	-0.459 (1.241)	-22.159 (86.346)	0.355 (3.359)
Dep. var. before interv.	1.036*** (0.002)	1.022*** (0.002)	1.018*** (0.001)	1.043*** (0.003)	1.038*** (0.003)	1.029*** (0.002)
Constant	9.205*** (1.521)	564.389*** (113.046)	29.975*** (3.696)	12.288*** (2.026)	779.288*** (140.985)	40.767*** (5.498)
Observations	5,572	5,572	5,572	5,572	5,572	5,572
R-squared	0.984	0.976	0.994	0.973	0.964	0.987

Note: Coefficient estimates from ordinary least squares estimations. * significant at 10%; ** sign. at 5%; *** sign. at 1%. The dependent variables report what learners achieved on the learning platform in the first four weeks after the first intervention email was sent in columns (1) - (3) and in the first eight weeks after the first intervention email was sent in columns (4) - (6). The dependent variable reports how many exercises a learner completed columns (1) and (4), how many points the learner collected in columns (2) and (5) and how much time the learner spent in columns (3) and (6). Values smaller than the 1th percentile or greater than the 99th percentile of the dependent variable are replaced by the value at the 1th or 99th percentile, respectively (winsorizing). (This is also done for the explanatory variable Dep.var. before interv.) The explanatory variable of main interest is identity treatment which takes the value of 1 if learner i participated in the identity treatment and is 0 otherwise (control group - reminder only). Days since registration measures the time a learner was registered on the platform before the first intervention email was sent. Active before intervention is a dummy variable that takes the value of 1 when a learner spent time on the platform before the intervention and 0 otherwise. The dummy variable Female takes the value of 1 when learner i is female and is 0 otherwise. Heteroscedasticity-robust Huber-White standard errors are in parentheses. Additionally, randomization inference based p-values, their standard error and the number of permutation to compute these p-values are reported for the identity treatment dummy at the bottom of the table. This takes into account the triple-wise stratification of the intervention.

Table A.8: Learning Behavior - Trimming (1th and 99th Percentile)

VARIABLES	(1)	(2)	(3)	(4)	(5)	(6)
	Exercises	After 4 weeks Points	Learn time	Exercises	After 8 weeks Points	Learn time
Identity treatment (d)	1.511* (0.838)	132.866** (63.402)	4.184** (2.082)	1.506 (1.062)	102.748 (74.400)	5.758* (3.103)
Months since registration	-0.636*** (0.091)	-40.665*** (6.910)	-2.424*** (0.227)	-0.761*** (0.116)	-47.537*** (8.114)	-3.338*** (0.338)
Active before treatment (d)	1.451* (0.853)	141.065** (64.468)	1.822 (2.137)	2.473** (1.080)	189.989** (75.643)	3.605 (3.185)
Female (d)	-0.123 (0.912)	-5.101 (68.953)	0.050 (2.263)	0.204 (1.155)	20.872 (80.923)	0.902 (3.372)
Dep. var. before interv.	1.018*** (0.002)	1.014*** (0.003)	1.017*** (0.001)	1.022*** (0.003)	1.018*** (0.003)	1.026*** (0.002)
Constant	8.252*** (1.487)	526.221*** (112.407)	30.063*** (3.691)	10.115*** (1.884)	656.189*** (131.935)	41.364*** (5.502)
Observations	5,517	5,517	5,517	5,517	5,517	5,517
R-squared	0.978	0.966	0.991	0.965	0.955	0.982

Note: Coefficient estimates from ordinary least squares estimations. * significant at 10%; ** sign. at 5%; *** sign. at 1%. The dependent variables report what learners achieved on the learning platform in the first four weeks after the first intervention email was sent in columns (1) - (3) and in the first eight weeks after the first intervention email was sent in columns (4) - (6). The dependent variable reports how many exercises a learner completed columns (1) and (4), how many points the learner collected in columns (2) and (5) and how much time the learner spent in columns (3) and (6). Values smaller than the 1th percentile or greater than the 99th percentile of the dependent variable are discarded (trimming). (This is also done for the explanatory variable Dep.var. before interv.) The explanatory variable of main interest is identity treatment which takes the value of 1 if learner i participated in the identity treatment and is 0 otherwise (control group - reminder only). Days since registration measures the time a learner was registered on the platform before the first intervention email was sent. Active before intervention is a dummy variable that takes the value of 1 when a learner spent time on the platform before the intervention and 0 otherwise. The dummy variable Female takes the value of 1 when learner i is female and is 0 otherwise. Heteroscedasticity-robust Huber-White standard errors are in parentheses. Additionally, randomization inference based p-values, their standard error and the number of permutation to compute these p-values are reported for the identity treatment dummy at the bottom of the table. This takes into account the triple-wise stratification of the intervention.

APPENDIX B - EXPERIMENTAL MATERIAL

On the following pages the Intervention Emails are displayed. The first email combines the 3 treatment-arms of the first mailing. The control group received the displayed email with the subject line "New German-learning application" (original: "Neue Deutschlern-App") excluding the green and red boxes. The identity gain treatment group's email subject was "New German-learning application: Become part of Germany" (original: "Neue Deutschlern-App: Werde Teil der deutschen Community") and its text included the text in the green boxes (without a background color). The identity loss treatment group's email subject was "New German-learning application: Stay part of Germany" (original: "Neue Deutschlern-App: Bleibe Teil der deutschen Community") and its text included the text in the red boxes (without a background color).

The other two emails were designed in a similar fashion with the control subject line of "Do not miss the new German-learning application" (original: "Verpassen Sie die neue Deutschlern-App nicht!") for the second mailing and "Last reminder: New German-learning application" (original: "Letzte Erinnerung: Neue Deutschlern-App").

Liebe/r Max Mustermann,

werden Sie Teil der deutschen Community: Deutsch lernen hilft, Teil der deutschen Gesellschaft zu werden. So finden Sie Anschluss und fühlen sich hier zuhause.

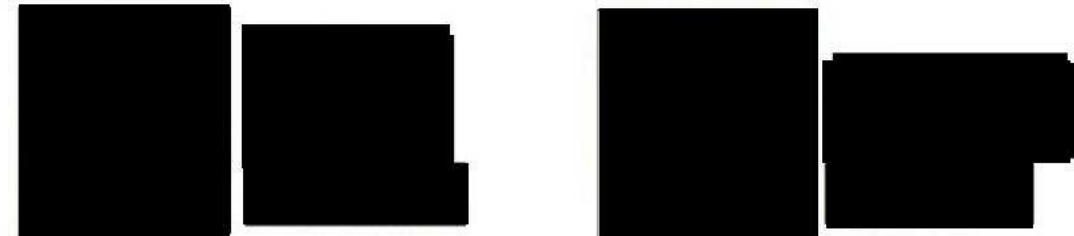
bleiben Sie Teil der deutschen Community: Deutsch lernen hilft, Teil der deutschen Gesellschaft zu bleiben. So verlieren Sie den Anschluss nicht und fühlen sich nicht fremd.

hat eine neue App zum Deutschlernen auf dem Smartphone. Für Sie ist die App und die Nutzung am Computer weitere 2 Monate kostenlos!

com im Web

App (Android*)

Viele Grüße



Dear Max Mustermann,

Become part of the German community: Learning German will help you become part of the German society. It will allow you to connect with others and help you feel at home.

Stay part of the German community: Learning German will help you stay part of the German society. It will ensure you stay connected and do not feel isolated.

has launched a new app for learning German. You can use the app as well as the desktop version free of charge for two more months.

on the Web

App (Android*)

Kind regards,

Your Team

Bei Problemen hier klicken. If you cannot view this e-mail correctly, please click here.

مرحبا بك

لكي تصبح جزءا من المجتمع الألماني: تعليم اللغة تسهل الاندماج في المجتمع و يربطك به و يعزز الانتماء اليه

لكي تبقي جزءا من المجتمع الألماني: تعليم اللغة يجعلك جزءا من هذا المجتمع ويقلل الإحساس بالغربة كما انه يحول دون فقدان حسن الارتباط بالمجتمع

أصدر "لينغوا تي فاو" تطبيق تعليم اللغة الألمانية الجديد للاستخدام على الهاتف الذكي. باستطاعتك ان تستخدمه على الهاتف والكمبيوتر مجانا لمدة شهرين

(Web)

App (Android*)

مع أطيب التحيات

Bei Problemen hier klicken. If you cannot view this e-mail correctly, please click here.

Max Mustermann محترم

بخشی از جامعه آلمان شوید: یاد زبان آلمانی کمک می کند برای این که شما بتوانید بخشی از جامعه آلمان شوید. به این ترتیب با مردم و فرهنگ آشنا شوید و در اینجا احساس راحتی داشته باشید

بخشی از جامعه آلمان بمانید: آموزش زبان آلمانی کمک می کند برای این که شما بتوانید بخشی از جامعه آلمان بمانید. به این ترتیب میتوانید با مردم در تماس باشید و در اینجا احساس بیگانگی نداشته باشید

شرکت "لينغوا تي وي" برنامه ای نو برای آموزش زبان در تلفن دارد. شما می توانید از این برنامه در تلفن و کمپوتر برای مدت دو ماه به صورت رایگان استفاده کنید

(Web)

App (Android*)

* Die App ist ab der Android-Version 5.1.2 verfügbar. Eine iOS-Version wird folgen. / The app is available from Android 5.1.2. There will be an iOS version in the future.

Impressum/ Imprint:



Abmeldelink | unsubscribe | إلغاء الاشتراك | لغو اشتراك

Liebe/r Max Mustermann,

nicht vergessen: Deutsch hilft Ihnen bei der **Jobsuche!**

Werden Sie Teil der deutschen Community: Deutsch lernen hilft, Teil der deutschen Gesellschaft zu werden. So finden Sie Anschluss und fühlen sich hier zuhause.

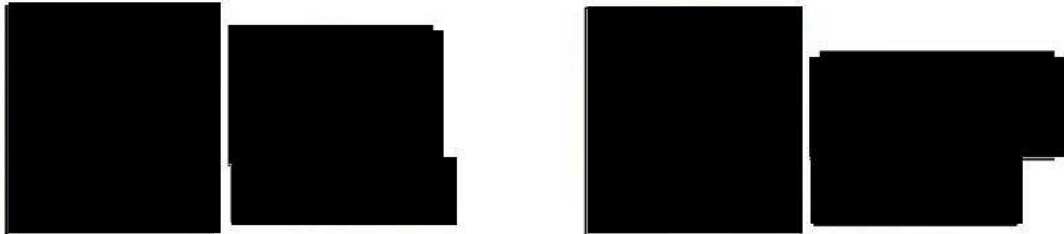
Werden Sie Teil der deutschen Community: Deutsch lernen hilft, Teil der deutschen Gesellschaft zu werden. So finden Sie Anschluss und fühlen sich hier zuhause.

hat eine neue App zum Deutschlernen auf dem Smartphone. Für Sie ist die App und die Nutzung am Computer weitere 2 Monate kostenlos!

com im Web

App (Android*)

Viele Grüße



Bei Problemen hier klicken. If you cannot view this e-mail correctly, please click here.

مرحبا بك

لا ينسى: تعليم اللغة الألمانية يساعدك في توحيد وظيفة

لكي تصبح جزءا من المجتمع الألماني: تعليم اللغة تسهل الاندماج في المجتمع و يربطك به و يعزز الانتماء اليه

لكي تبقي جزءا من المجتمع الألماني: تعليم اللغة يجعلك جزءا من هذا المجتمع ويقلل الإحساس بالغربة كما انه يحول دون فقدان حسن الارتباط بالمجتمع

أصدر "لينكوا تي فاو" تطبيق تعليم اللغة الألمانية الجديد للاستخدام على الهاتف الذكي. باستطاعتك ان تستخدمه على الهاتف والكمبيوتر مجانا

(Web)

App (Android*)

مع أطيب التحيات

Bei Problemen hier klicken. If you cannot view this e-mail correctly, please click here.

محترم Max Mustermann

فراموش نکنید: زبان آلمانی برای پیدا کردن کار کمک می کند

بخشی از جامعه آلمان شوید: یاد زبان آلمانی کمک می کند برای این که شما بتوانید بخشی از جامعه آلمان شوید. به این ترتیب با مردم و فرهنگ آشنا شوید و در اینجا احساس راحتی داشته باشید

بخشی از جامعه آلمان بمانید: آموزش زبان آلمانی کمک می کند برای این که شما بتوانید بخشی از جامعه آلمان بمانید. به این ترتیب می توانید با مردم در تماس باشید و در اینجا احساس بیگانگی نداشته باشید

شرکت "لینگوا تی وی" برنامه ای نو برای آموزش زبان در تلفن دارد. شما می توانید از این برنامه در تلفن و کامپیوتر برای مدت دو ماه به صورت رایگان استفاده کنید

(Web)

App (Android*)

با احترام

Team

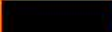
Dear Max Mustermann,

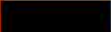
Don't forget: German skills help finding a job.

Become part of the German community: Learning German will help you become part of the German society. It will allow you to connect with others and help you feel at home.

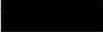
Become part of the German community: Learning German will help you become part of the German society. It will allow you to connect with others and help you feel at home.

has launched a new app for learning German. You can use the app as well as the desktop version free of charge for two more months.

 on the Web

 App (Android*)

Kind regards,

Your  Team

* Die App ist ab der Android-Version 5.1.2 verfügbar. Eine iOS-Version wird folgen. / The app is available from Android 5.1.2. There will be an iOS version in the future.

 Impressum/ Imprint:









Abmeldelink | unsubscribe | إلغاء الاشتراك | لغو اشتراك

Liebe/r Max Mustermann,

Deutsch ist wichtig! Wir laden Sie ein, weiterhin **kostenlos** Deutsch zu lernen.

Werden Sie Teil der deutschen Community: Deutsch lernen hilft, Teil der deutschen Gesellschaft zu werden. So finden Sie Anschluss und fühlen sich hier zuhause.

Werden Sie Teil der deutschen Community: Deutsch lernen hilft, Teil der deutschen Gesellschaft zu werden. So finden Sie Anschluss und fühlen sich hier zuhause.

hat eine neue App zum Deutschlernen auf dem Smartphone. Für Sie ist die App und die Nutzung am Computer weitere 2 Monate kostenlos!

com im Web

App (Android*)

Wenn Ihnen die App gefällt, bitte mit 5 Sternen bewerten ... und empfehle sie weiter!

Viele Grüße



Bei Problemen hier klicken. If you cannot view this e-mail correctly, please click here.

مرحبا بك

اللغة الألمانية مهمة: نرجو أن تستمر بتعليم اللغة مجاناً

لكي تصبح جزءاً من المجتمع الألماني: تعليم اللغة تسهل الاندماج في المجتمع و يربطك به و يعزز الانتماء اليه

لكي تبقي جزءاً من المجتمع الألماني: تعليم اللغة يجعلك جزءاً من هذا المجتمع ويقلل الإحساس بالغربة كما انه يحول دون فقدان حسن الارتباط

أصدر "لينكوا تي فاو" تطبيق تعليم اللغة الألمانية الجديد للاستخدام على الهاتف الذكي. باستطاعتك ان تستخدمه على الهاتف والكمبيوتر مجاناً لمدة شهرين

(Web)

App (Android*)

إذ عجبك هذه التطبيق لتعليم اللغة من فضلك أعطيه خمسة نجم و أبلغ أصدقائك عنه

مع أطيب التحيات

Bei Problemen hier klicken. If you cannot view this e-mail correctly, please click here.

محترم Max Mustermann

زبان آلمانی مهم است! ما شما را دعوت می کنیم که رایگان زبان آلمانی را یاد بگیرید

بخشی از جامعه آلمان شوید: یاد زبان آلمانی کمک می کند برای این که شما بتوانید بخشی از جامعه آلمان شوید. به این ترتیب با مردم و فرهنگ آشنا شوید و در اینجا احساس راحتی داشته باشید

بخشی از جامعه آلمان بمانید: آموزش زبان آلمانی کمک می کند برای این که شما بتوانید بخشی از جامعه آلمان بمانید. به این ترتیب میتوانید با مردم در تماس باشید و در اینجا احساس بیگانگی نداشته باشید

شرکت "لینگوا تی وی" برنامه ای نو برای آموزش زبان در تلفن دارد. شما می توانید از این برنامه در تلفن و کمپوتر برای مدت دو ماه به صورت رایگان استفاده کنید

(Web)

App (Android*)

اگر از برنامه لینگوا تی وی خوشتان آمده، آن را با ۵ ستاره ارزیابی کنید و به دوستان شما از آن تعریف کنید

با احترام

Team

Dear Max Mustermann,

Learning German is important! We invite you to continue learning German for free.

Become part of the German community: Learning German will help you become part of the German society. It will allow you to connect with others and help you feel at home.

Become part of the German community: Learning German will help you become part of the German society. It will allow you to connect with others and help you feel at home.

has launched a new app for learning German. You can use the app as well as the desktop version free of charge for two more months.

on the Web

App (Android*)

If you like the app, please rate it with 5 stars and recommend it!

Kind regards,

Your Team

* Die App ist ab der Android-Version 5.1.2 verfügbar. Eine iOS-Version wird folgen. / The app is available from Android 5.1.2. There will be an iOS version in the future.

Impressum/ Imprint:



[Abmeldelink](#) | [unsubscribe](#) | [لغو الاشتراك](#) | [إلغاء الاشتراك](#)

References

- Akerlof, G. A. and Kranton, R. E. (2000). Economics and Identity. *Quarterly Journal of Economics*, CXV(3):715–753.
- Akerlof, R. (2017). Value Formation : The Role of Esteem. *Games and Economic Behavior*, 102:1–19.
- Austen-Smith, D. and Fryer, R. (2005). An economic analysis of “acting white”. *Quarterly Journal of Economics*, 120(May):551–583.
- Banks, S., Salovey, P., Greener, S., Rothman, A. J., Moyer, A., Beauvais, J., and Epel, E. (1995). The effects of message framing on mammography utilization.
- Battu, H. and Zenou, Y. (2009). Oppositional identities and employment for ethnic minorities: Evidence from England. *Economic Journal*, 120(542):52–71.
- Bauer, T., Epstein, G. S., and Gang, I. N. (2005). Enclaves, language, and the location choice of migrants. *Journal of Population Economics*, 18(4):649–662.
- Benabou, R. and Tirole, J. (2011). Identity, Morals and Taboos. *Quarterly Journal of Economics*, 126(2):805–855.
- Benjamin, D. J., Choi, J. J., and Fisher, G. (2016). Religious Identity and Economic Behavior. *Review of Economics and Statistics*, 98(4):617–637.
- Benjamin, D. J., Choi, J. J., and Strickland, a. J. (2010). Social Identity and Preferences and Preferences Identity. *American Economic Association*, 100(4):1913–1928.
- Bernhard, B. H., Fehr, E., and Fischbacher, U. R. S. (2016). American Economic Association Group Affiliation and Altruistic Norm Enforcement Author (s): Helen Bernhard , Ernst Fehr and Urs Fischbacher Source : The American Economic Review , Vol . 96 , No . 2 (May , 2006) , pp . 217-221 Published by : American Ec. *American Economic Review*, 96(2):217–221.

- Bertrand, M., Karlan, D., Mullainathan, S., Shafir, E., and Zinman, J. (2010). What’s Advertising Content Worth? Evidence from a Consumer Credit Marketing Field Experiment. *Quarterly Journal of Economics*, 125(1):263–305.
- Bisin, A., Patacchini, E., Verdier, T., and Zenou, Y. (2011). Formation and persistence of oppositional identities. *European Economic Review*, 55(8):1046–1071.
- Bisin, A., Patacchini, E., Verdier, T., and Zenou, Y. (2016). Bend it like Beckham: Ethnic identity and integration. *European Economic Review*, 90:146–164.
- Bisin, A. and Verdier, T. (2000). ”Beyond the Melting Pot”: Cultural Transmission, Marriage, and the Evolution of Ethnic and Religious Traits. *Quarterly Journal of Economics*, 115(3):955–988.
- Brücker, H., Rother, N., and Schupp, J. (2016). IAB-BAMF-SOEP-Befragung von Geflüchteten: Überblick und erste Ergebnisse. Technical report, IAB.
- Brücker, H., Rother, N., and Schupp, J. (2017). IAB-BAMF-SOEP-Befragung von Geflüchteten: Überblick und erste Ergebnisse. Technical report, DIW Berlin.
- Brücker, H., Rother, N., and Schupp, J. (2018). IAB-BAMF-SOEP- Befragung von Geflüchteten 2016: Studiendesign, Feldergebnisse sowie Analysen zu schulischer wie beruflicher Qualifikation, Sprachkenntnissen sowie kognitiven Potenzialen. Technical report, Bundesamt für Migration und Flüchtlinge.
- Calzolari, G. and Nardotto, M. (2016). Effective Reminders. *Management Science*, 51(7):1 – 18.
- Cameron, L., Erkal, N., Gangadharan, L., and Zhang, M. (2015). Cultural integration: Experimental evidence of convergence in immigrants’ preferences. *Journal of Economic Behavior and Organization*, 111:38–58.
- Candelo, N., Croson, R. T., and Li, S. X. (2017). Identity and social exclusion: an experiment with Hispanic immigrants in the U.S. *Experimental Economics*, 20(2):460–480.

- Charness, G., Cobo-Reyes, R., and Jiménez, N. (2014). Identities, selection, and contributions in a public-goods game. *Games and Economic Behavior*, 87:322–338.
- Chen, Y., Farzan, R., Kraut, R., YeckehZaare, I., and Zhang, A. F. (2017). ExpertIdeas: Incentivizing Domain Experts to Contribute to Wikipedia.
- Chen, Y. and Li, S. X. (2009). Group Identity and Social Preferences. *The American Economic Review*, 99(1):431–457.
- Chen, Y., Li, S. X., Liu, T. X., and Shih, M. (2014). Which hat to wear? Impact of natural identities on coordination and cooperation. *Games and Economic Behavior*, 84:58–86.
- Chiswick, B. R. and Miller, P. W. (1995). The Endogeneity between Language and Earnings: International Analyses. *Journal of Labor Economics*, 13(2):246–288.
- Clots-Figueras, I. and Masella, P. (2013). Education, language and identity. *Economic Journal*, 123(570):332–357.
- Cohn, A., Fehr, E., and Marechal, M. A. (2014). Business culture and dishonesty in the banking industry. *Nature*, 516(729):86–89.
- Cohn, A., Maréchal, M. A., and Noll, T. (2010). Bad boys: How criminal identity salience affects rule violation. *Review of Economic Studies*, 82(4):1289–1308.
- Connolly, K. (2015). Germany greets refugees with help and kindness at Munich central station.
- Constant, A. F., Gataullina, L., and Zimmermann, K. F. (2009). Ethnosizing immigrants. *Journal of Economic Behavior and Organization*, 69(3):274–287.
- Damgaard, M. T. and Nielsen, H. S. (2017). The use of nudges and other behavioural approaches in education.
- Davis, J. B. (2007). Akerlof and Kranton on identity in economics: Inverting the analysis. *Cambridge Journal of Economics*, 31(3):349–362.

- Dehdari, S. and Gehring, K. (2017). The Origins of Common Identity: Division, Homogenization Policies and Identity Formation in Alsace-Lorraine.
- Dustmann, C. (1994). Speaking fluency, writing fluency and earnings of migrants. *Journal of Population Economics*, 7(2):133–156.
- Dustmann, C., Fabbri, F., Preston, I., and Wadsworth, J. (2003). Labour market performance of immigrants in the UK labour market. *Economic Journal, Royal Economic Society*, 113:695–717.
- Epstein, G. S. and Gang, I. N. (2010). Chapter 1 Migration and Culture. *Migration and Culture*, 8(5123):1–21.
- Epstein, G. S. and Heizler, O. (2015). Ethnic identity: a theoretical framework. *IZA Journal of Migration*, 4(1):9.
- Escueta, M., Quan, V., Nickow, A. J., and Oreopoulos, P. (2017). Education Technology: An Evidence-based Review. *NBER Working Paper Series*.
- Falk, A., Becker, A., Dohmen, T., Enke, B., Huffman, D. B., and Sunde, U. (2017). Global Evidence on Economic Preferences.
- Fordham, S. and Ogbu, J. (1986). Black Students School Success: Coping with the Burden of Acting White (Annotated). *The urban review*, 18(3):176–206.
- Fryer, R. G., Levitt, S., List, J., and Sadoff, S. (2012). Enhancing the Efficacy of Teacher Incentives through Loss Aversion: A Field Experiment.
- Fryer, R. G. and Torelli, P. (2010). An empirical analysis of 'acting white'. *Journal of Public Economics*, 94(5-6):380–396.
- Gallagher, K. M. and Updegraff, J. A. (2012). Health message framing effects on attitudes, intentions, and behavior: A meta-analytic review. *Annals of Behavioral Medicine*, 43(1):101–116.

- Georgiadis, A. and Manning, A. (2013). One nation under a groove? Understanding national identity. *Journal of Economic Behavior and Organization*, 93:166–185.
- Goette, L., Huffman, D., and Meier, S. (2006). The impact of group membership on cooperation and norm enforcement. *American Economic Review*, 96(2):212–216.
- Grote, N., Klausmann, T., and Scharfbillig, M. (2017). The Absolute and Relative Effect of an Identity Gain and Loss Treatment on Learning Behavior.
- Hansen, P. G. (2016). The Definition of Nudge and Libertarian Paternalism: Does the Hand Fit the Glove? *European Journal of Risk Regulation*, 7(1):155–174.
- Heap, S. P. H. and Zizzo, D. J. (2009). The value of groups. *American Economic Review*, 99(1):295–323.
- Heß, S. (2017). Randomization inference with Stata: A guide and software. *The Stata Journal*, 17(3):630–651.
- Hett, F., Kröll, M., and Mechtel, M. (2017). Choosing Who You Are: The Structure and Behavioral Effects of Revealed Identification Preferences.
- Hicks, D. L., Santacreu-Vasut, E., and Shoham, A. (2015). Does mother tongue make for women’s work? Linguistics, household labor, and gender identity. *Journal of Economic Behavior and Organization*, 110:19–44.
- Hoff, K. and Pandey, P. (2006). Discrimination, Social Identity, and Durable Inequalities. *American Economic Review*, 96(2):206–211.
- Hoff, K. and Pandey, P. (2014). Making up people-The effect of identity on performance in a modernizing society. *Journal of Development Economics*, 106:118–131.
- Hossain, T. and List, J. A. (2012). The Behavioralist Visits the Factory: Increasing Productivity Using Simple Framing Manipulations. *Management Science*, 58(12):2151–2167.
- Jenkins, R. (2014). *Social Identity*. Routledge.

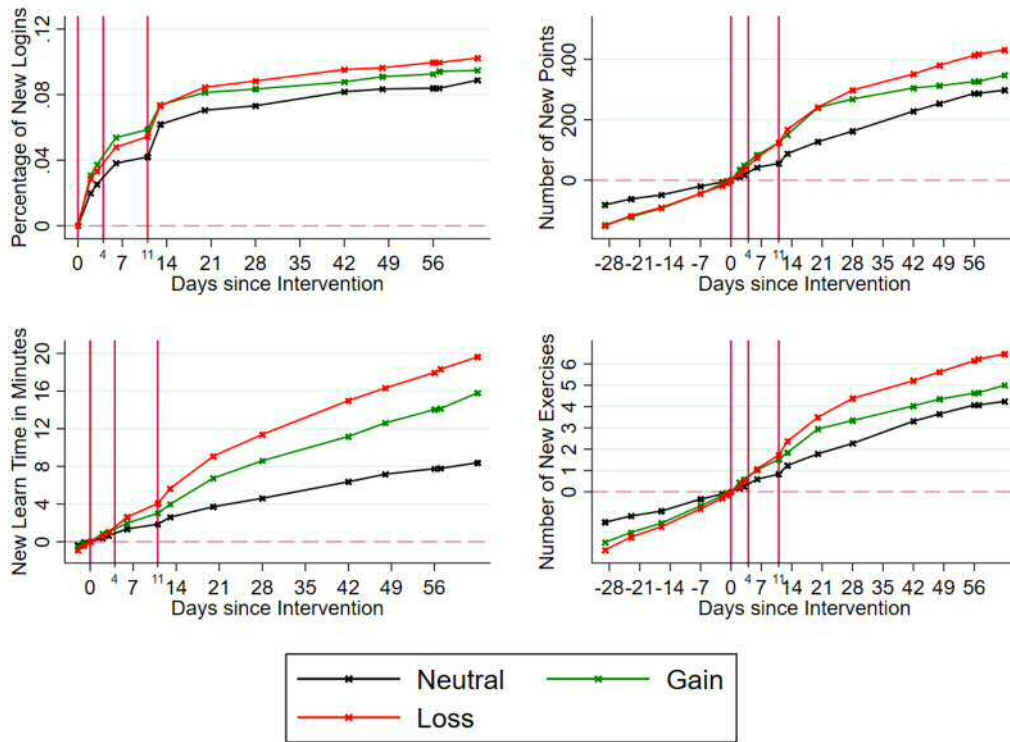
- Kahneman Daniel and Tversky Amos (2011). An Analysis of Decision under Risk. *Econometrica*, Vol. 47, No. 2, March 1979, 47(2):263–292.
- Karlan, D., McConnell, M., Mullainathan, S., and Zinman, J. (2016). Getting to the Top of Mind: How Reminders Increase Saving. *Management Science*, 62(12):3393–3411.
- Karlan, D., Morten, M., and Zinman, J. (2012). A personal touch: Text messaging for loan repayment.
- Keysar, B., Hayakawa, S. L., and An, S. G. (2012). The Foreign-Language Effect. *Psychological Science*, 23(6):661–668.
- Kranton, R., Pease, M., Sanders, S., and Huettel, S. (2016). Groupy and Non-Groupy Behavior: Deconstructing Bias in Social Preferences. (May).
- Lazear, E. (1999). Culture and Language. *Journal of Political Economy*, 107(S6):S95–S126.
- Leary, M. R., Wheeler, D. S., and Jenkins, T. B. (1986). Aspects of Identity and Behavioral Preference: Studies of Occupational and Recreational Choice. *Social Psychology Quarterly*, 49(1):11–18.
- Levitt, S. D., List, J. A., Neckermann, S., and Sadoff, S. (2016). The behavioralist goes to school: Leveraging behavioral economics to improve educational performance. *American Economic Journal: Economic Policy*, 8(4):183–219.
- MacKinnon, J. G. and White, H. (1985). Some heteroskedasticity-consistent covariance matrix estimators with improved finite sample properties. *Journal of Econometrics*, 29(3):305–325.
- Manning, A. and Roy, S. (2009). Culture clash or culture club? National identity in Britain. *Economic Journal*, 120(542):72–100.
- McKenzie, D. (2012). Beyond baseline and follow-up: The case for more T in experiments. *Journal of Development Economics*, 99(2):210–221.

- Mcmanus, W., Gould, W., and Welch, F. (1983). Earnings of Hispanic Men: The Role of English Language Proficiency. *Journal of Labor Economics*, 1(2):101–130.
- Monscheuer, O. (2018). National Attachment and the Integration of Second Generation Immigrants in the U.S.
- Morrisson, C. and Murtin, F. (2009). The Century of Education. *Journal of Human Capital*, 3(1):1–42.
- Ogbu, J. U. (1999). Beyond Language: Ebonics, Proper English, and Identity in a Black-American Speech Community. *American Educational Research Journal*, 36(2):147–184.
- O’Keefe, D. J. and Jensen, J. D. (2009). The relative persuasiveness of gain-framed and loss-framed messages for encouraging disease detection behaviors: A meta-analytic review. *Journal of Communication*, 59(2):296–316.
- Patacchini, E. and Zenou, Y. (2016). Racial identity and education in social networks. *Social Networks*, 44:85–94.
- Peirce, B. N. (1995). Social Identity, Investment, and Language Learning. *TESOL Quarterly*, 29(1):9.
- Stryker, S. and Serpe, R. T. (1982). Commitment, identity salience, and role behavior: Theory and research example. In *Personality, roles, and social behavior*, pages 199–218. Springer.
- Tajfel, H. (1974). Social identity and intergroup behaviour. *Social Science Information*, 13(2):65–93.
- Tajfel, H. (1978). *Differentiation between social groups: Studies in the social psychology of intergroup relations*. Academic Press, Oxford, England.
- Tajfel, H. and Turner, J. C. (1979). An Integrative Theory of Intergroup Conflict. In Austin, W. and Worchel, S., editors, *The social psychology of intergroup relations*, pages 33–47. Brooks/Cole, Monterey, CA.

- Tajfel, H. and Turner, J. C. (1986). The Social Identity Theory of Intergroup Behavior. In Worchel, S. and Austin, W. G., editors, *The social psychology of intergroup relations*, pages 7–24. Nelson-Hall, Chicago, US.
- Thaler, R. H. (1980). Towards a positive theory of consumer choice. *Journal of Economic Behavior & Organization*, 1:39–60.
- Thaler, R. H. and Sunstein, C. R. (2008). *Nudge: Improving decisions about health, wealth, and happiness*. Yale University Press.
- Tversky, A. and Kahneman, D. (1991). Loss Aversion in Riskless Choice: A Reference-Dependent Model. *The Quarterly Journal of Economics*, 106(4):1039–1061.

FIGURES AND TABLES

Figure I: Learning Behavior Over Time



Note: The figure reports the aggregate Learning behavior over the whole intervention period. Percentage of New Logins is the fraction of learners who logged onto the learning platform after the intervention (left top), New Points is the number of points the learner collected after the first mail (top right), New Learn Time is how much time the learner spent on the platform after the first mail (bottom left) and New Exercises reports how many exercises a learner completed after the first mail (bottom right). Red vertical lines mark dates on which we send emails to all participants.

Table I: Descriptives

	N	Mean	Median	SD	Min	Max
<i>Control variables</i>						
Months since registration	5573	13.176	14	4.344	2	24
Active before treatment (d)	5573	0.549	1	0.498	0	1
Female (d)	5572	0.256	0	0.436	0	1
<i>Mailings</i>						
Opened (d)	5570	0.660	1	0.474	0	1
Clicked (d)	5434	0.224	0	0.417	0	1
<i>Learning behavior before treatment</i>						
Logins	5573	20.439	6	43.373	1	1125
Exercises	5573	94.953	1	269.275	0	3481
Points	5573	5755.625	72	16197.524	0	215057
Learn time	5573	405.410	12	1142.477	0	15517

Note: Months since registration measures the time a learner was registered on the platform before the first intervention email was sent. Active before intervention is a dummy variable that takes the value of 1 when a learner spent time on the platform before the intervention and 0 otherwise. The dummy variable Female takes the value of 1 when learner i is female and is 0 otherwise. Opened is a dummy variable that takes the value of 1 when the learner opened at least one of the intervention emails. Clicked is a dummy equal to 1 if the learner clicked at least once on a link to the language-learning platform presented in the intervention emails. Logins is the number of times the learner logged onto the learning platform, Exercises reports how many exercises a learner completed, Points how many points the learner collected and Learn time how much time the learner spent, all measured by the platform prior to the intervention.

Table II: Reactions to Mailings

VARIABLES	(1) Opened (d)	(2) Clicked (d)	(3) New login 4 weeks (d)
Identity treatment (d)	0.081*** (0.014)	0.023* (0.012)	0.013* (0.007)
Months since registration	-0.001 (0.001)	0.000 (0.001)	-0.009*** (0.001)
Active before treatment (d)	0.038*** (0.013)	0.025** (0.011)	0.079*** (0.007)
Female (d)	-0.048*** (0.015)	-0.060*** (0.012)	-0.020** (0.008)
Constant	0.608*** (0.024)	0.210*** (0.021)	0.147*** (0.015)
Observations	5,569	5,433	5,572
R-squared	0.010	0.005	0.041
RI p-value of ident. treat.	0	0.046	0.067
RI SE of p-value	0	0.0021	0.0025
RI repetitions	10000	10000	10000

Note: The dependent variable is a dummy variable that takes the value of 1 when the learner opened at least one of the intervention emails in column (1), when the learner clicked at least once on a link to the language-learning platform presented in the intervention emails in column (2) and when the learner logged onto the learning platform at least once within four weeks of the intervention in column (3). The explanatory variable of main interest is identity treatment which takes the value of 1 if learner i participated in the identity treatment and is 0 otherwise (control group - reminder only). Months since registration measures the time a learner was registered on the platform before the first intervention email was sent. Active before intervention is a dummy variable that takes the value of 1 when a learner spent time on the platform before the intervention and 0 otherwise. The dummy variable Female takes the value of 1 when learner i is female and is 0 otherwise. Additionally, randomization inference based p-values, their standard error and the number of permutation to compute these p-values are reported for the identity treatment dummy at the bottom of the table. This takes into account the triple-wise stratification of the intervention. Heteroscedasticity-robust Huber-White standard errors are in parentheses. Coefficient estimates from ordinary least squares estimations.

* $p < 0.1$, ** $p < 0.05$, *** $p < 0.01$. Standard errors in parentheses.

Table III: Learning Behavior

VARIABLES	(1)	(2)	(3)	(4)	(5)	(6)
	Exercises	After 4 weeks Points	Learn time	Exercises	After 8 weeks Points	Learn time
Identity treatment (d)	1.571** (0.753)	119.281** (54.931)	5.390*** (1.811)	1.294 (1.190)	79.907 (78.872)	8.265*** (2.867)
Months since registration	-0.762*** (0.164)	-48.309*** (10.664)	-2.466*** (0.491)	-0.966*** (0.208)	-60.295*** (13.571)	-3.368*** (0.664)
Active before treatment (d)	1.982*** (0.693)	145.512** (60.112)	5.657*** (1.416)	3.211*** (0.977)	212.082*** (73.946)	8.091*** (2.692)
Female (d)	-0.318 (1.002)	-22.091 (77.493)	-0.580 (2.510)	-0.875 (1.321)	-45.514 (95.991)	-1.039 (4.010)
Dep. var. before interv.	1.015*** (0.003)	1.015*** (0.004)	1.010*** (0.002)	1.022*** (0.005)	1.021*** (0.005)	1.018*** (0.005)
Constant	9.939*** (2.117)	641.055*** (139.802)	29.971*** (6.171)	13.238*** (2.705)	856.849*** (175.723)	40.540*** (8.195)
Observations	5,572	5,572	5,572	5,572	5,572	5,572
R-squared	0.988	0.980	0.995	0.976	0.969	0.989
RI p-value of ident. treat.	0.062	0.057	0.0081	0.27	0.33	0.011
RI SE of p-value	0.0024	0.0023	0.00090	0.0045	0.0047	0.0010
RI repetitions	10000	10000	10000	10000	10000	10000

Note: The dependent variables report what learners achieved on the learning platform in the first four weeks after the first intervention email was sent in columns (1) - (3) and in the first eight weeks after the first intervention email was sent in columns (4) - (6). The dependent variable reports how many exercises a learner completed columns (1) and (4), how many points the learner collected in columns (2) and (5) and how much time the learner spent in columns (3) and (6). The explanatory variable of main interest is identity treatment which takes the value of 1 if learner *i* participated in the identity treatment and is 0 otherwise (control group - reminder only). Days since registration measures the time a learner was registered on the platform before the first intervention email was sent. Active before intervention is a dummy variable that takes the value of 1 when a learner spent time on the platform before the intervention and 0 otherwise. The dummy variable Female takes the value of 1 when learner *i* is female and is 0 otherwise. Additionally, randomization inference based p-values, their standard error and the number of permutation to compute these p-values are reported for the identity treatment dummy at the bottom of the table. This takes into account the triple-wise stratification of the intervention. Heteroscedasticity-robust Huber-White standard errors are in parentheses. Coefficient estimates from ordinary least squares estimations.

* $p < 0.1$, ** $p < 0.05$, *** $p < 0.01$. Standard errors in parentheses.

Table IV: Effect of Gain and Loss Treatment - Mailing and Learning Outcomes

VARIABLES	(1)	(2)	(3)	(4)	(5)	(6)
	Opened (d)	Clicked (d)	New login 4 weeks (d)	Exercises	After 4 weeks Points	Learn time
Ident. treat. = 1, Gain	0.082*** (0.016)	0.023 (0.014)	0.011 (0.009)	1.128 (0.852)	108.285 (72.345)	4.155** (1.917)
Ident. treat. = 2, Loss	0.081*** (0.016)	0.023* (0.014)	0.016* (0.009)	2.013** (1.015)	130.282* (67.330)	6.626** (2.726)
Months since registration	-0.001 (0.001)	0.000 (0.001)	-0.009*** (0.001)	-0.762*** (0.164)	-48.310*** (10.665)	-2.466*** (0.491)
Active before treatment (d)	0.038*** (0.013)	0.025** (0.011)	0.079*** (0.007)	1.986*** (0.693)	145.590** (60.028)	5.665*** (1.419)
Female (d)	-0.048*** (0.015)	-0.060*** (0.012)	-0.020** (0.008)	-0.316 (1.002)	-22.035 (77.452)	-0.575 (2.510)
Dep. var. before interv.				1.015*** (0.003)	1.015*** (0.004)	1.010*** (0.002)
Constant	0.608*** (0.024)	0.210*** (0.021)	0.147*** (0.015)	9.939*** (2.117)	641.064*** (139.818)	29.970*** (6.169)
Observations	5,569	5,433	5,572	5,572	5,572	5,572
R-squared	0.010	0.005	0.041	0.988	0.980	0.995
RI p-value of gain treat.	0	0.21	0.32	0.39	0.27	0.20
RI SE of p-value of gain treat.	0	0.0041	0.0047	0.0049	0.0044	0.0040
RI p-value of gain loss.	0	0.20	0.17	0.12	0.18	0.040
RI SE of p-value of loss treat.	0	0.0040	0.0037	0.0033	0.0038	0.0020
RI repetitions	10000	10000	10000	10000	10000	10000

Note: Coefficient estimates from ordinary least squares estimations. * significant at 10%; ** sign. at 5%; *** sign. at 1%. The dependent variable is a dummy variable that takes the value of 1 when the learner opened at least one of the intervention emails in column (1), when the learner clicked at least once on a link to the language-learning platform presented in the intervention emails in column (2) and when the learner logged onto the learning platform at least once within four weeks of the intervention in column (3). The dependent variables report what learners achieved on the learning platform in the first four weeks after the first intervention email was sent in columns (4) - (6). The dependent variable reports how many exercises a learner completed columns (4), how many points the learner collected in columns (5) and how much time the learner spent in columns (6). The explanatory variables of main interest are the two versions of the identity treatment. Both the identity gain and identity loss treatment variables take the value of 1 if learner i participated in that specific treatment and 0 otherwise (control group - reminder only). Days since registration measures the time a learner was registered on the platform before the first intervention email was sent. Active before intervention is a dummy variable that takes the value of 1 when a learner spent time on the platform before the intervention and 0 otherwise. The dummy variable Female takes the value of 1 when learner i is female and is 0 otherwise. Heteroscedasticity-robust Huber-White standard errors are in parentheses. Additionally, randomization inference based p-values, their standard error and the number of permutation to compute these p-values are reported for the treatment dummies at the bottom of the table. This takes into account the triple-wise stratification of the intervention.

Table V: Interaction with Prior Activity Learning

VARIABLES	(1)	(2)	(3)	(4)	(5)	(6)
	Opened (d)	Clicked (d)	New login 4 weeks (d)	Exercises	After 4 weeks Points	Learn time
Identity treatment (d)	0.095*** (0.021)	0.028* (0.017)	0.004 (0.008)	0.255 (0.379)	23.306 (35.363)	-0.183 (0.718)
Identity treatment * Active before treatment (d)	-0.025 (0.027)	-0.010 (0.024)	0.017 (0.014)	2.401* (1.388)	175.135* (102.000)	10.171*** (3.326)
Months since registration	-0.001 (0.001)	0.000 (0.001)	-0.009*** (0.001)	-0.762*** (0.164)	-48.320*** (10.662)	-2.467*** (0.491)
Active before treatment (d)	0.055** (0.023)	0.031* (0.019)	0.068*** (0.011)	0.383 (0.933)	28.920 (60.245)	-1.121 (2.280)
Female (d)	-0.047*** (0.015)	-0.060*** (0.012)	-0.021*** (0.008)	-0.350 (1.000)	-24.361 (77.272)	-0.713 (2.520)
Dep. var. before interv.				1.015*** (0.003)	1.015*** (0.004)	1.010*** (0.002)
Constant	0.599*** (0.026)	0.207*** (0.022)	0.153*** (0.015)	10.825*** (2.262)	705.700*** (148.609)	33.726*** (6.777)
Observations	5,569	5,433	5,572	5,572	5,572	5,572
R-squared	0.010	0.005	0.041	0.988	0.981	0.995

Note: Coefficient estimates from ordinary least squares estimations. * significant at 10%; ** sign. at 5%; *** sign. at 1%. The dependent variable is a dummy variable that takes the value of 1 when the learner opened at least one of the intervention emails in column (1), when the learner clicked at least once on a link to the language-learning platform presented in the intervention emails in column (2) and when the learner logged onto the learning platform at least once within four weeks of the intervention in column (3). The dependent variables report what learners achieved on the learning platform in the first four weeks after the first intervention email was sent in columns (4) - (6). The explanatory variable of main interest is the identity treatment interacted with Activity before treatment which takes the value of 1 if learner i participated in the identity treatment and was active before our intervention and is 0 otherwise. Days since registration measures the time a learner was registered on the platform before the first intervention email was sent. Active before intervention is a dummy variable that takes the value of 1 when a learner spent time on the platform before the intervention and 0 otherwise. The dummy variable Female takes the value of 1 when learner i is female and is 0 otherwise. Heteroscedasticity-robust Huber-White standard errors are in parentheses.

Table VI: Interactions with Clicked on German Link

VARIABLES	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
	New login (d)	After 4 weeks Exercises	Points	Learn time	New login (d)	After 8 weeks Exercises	Points	Learn time
Identity treatment (d)	-0.002 (0.007)	1.169 (0.751)	70.399 (54.131)	3.681* (1.922)	0.000 (0.007)	0.971 (1.087)	58.439 (70.067)	4.348 (2.678)
Clicked (d)	0.062** (0.026)	6.561 (4.019)	481.180* (251.224)	2.478 (3.732)	0.085*** (0.028)	11.954 (7.598)	900.718* (469.907)	8.570 (8.808)
Identity treatment * Clicked (d)	0.053 (0.033)	-5.366 (4.229)	-325.688 (276.967)	-3.627 (4.966)	0.044 (0.035)	-11.307 (7.745)	-769.364 (486.634)	-9.664 (10.087)
Identity * Clicked German Link (d)	0.003 (0.046)	12.106** (4.866)	917.012*** (342.149)	19.595** (8.756)	0.002 (0.048)	21.295** (8.663)	1,431.376** (571.798)	46.794*** (16.595)
Clicked German Link (d)	0.097*** (0.037)	-5.935 (4.144)	-440.054* (261.335)	-1.414 (4.402)	0.081** (0.039)	-10.116 (7.878)	-691.784 (507.034)	-10.892 (9.182)
Months since registration	-0.009*** (0.001)	-0.779*** (0.168)	-49.362*** (10.908)	-2.520*** (0.502)	-0.010*** (0.001)	-0.988*** (0.212)	-61.653*** (13.871)	-3.439*** (0.677)
Active before treatment (d)	0.076*** (0.007)	1.996*** (0.695)	145.603** (60.362)	5.850*** (1.468)	0.083*** (0.007)	3.197*** (0.971)	209.898*** (73.982)	8.288*** (2.767)
Female (d)	-0.014* (0.008)	-0.175 (1.037)	-7.807 (81.039)	-0.476 (2.553)	-0.016* (0.008)	-0.658 (1.352)	-25.115 (99.415)	-0.714 (4.017)
Dep. var. before interv.		1.014*** (0.003)	1.014*** (0.004)	1.010*** (0.002)		1.021*** (0.005)	1.020*** (0.005)	1.018*** (0.005)
Constant	0.127*** (0.015)	9.508*** (2.162)	608.301*** (144.750)	30.324*** (6.380)	0.147*** (0.016)	12.300*** (2.681)	773.525*** (178.681)	41.064*** (8.428)
Observations	5,433	5,433	5,433	5,433	5,433	5,433	5,433	5,433
R-squared	0.105	0.988	0.981	0.995	0.105	0.976	0.969	0.989

Note: Coefficient estimates from ordinary least squares estimations. * significant at 10%; ** sign. at 5%; *** sign. at 1%. The dependent variables report what learners achieved on the learning platform in the first four weeks after the first intervention email was sent in columns (1) - (4) and in the first eight weeks after the first intervention email was sent in columns (5) - (8). The dependent variable is a dummy variable that takes the value of 1 when the learner logged in at least once in column (1) and (5), how many exercises a learner completed columns (2) and (6), how many points the learner collected in columns (3) and (7) and how much time the learner spent in columns (4) and (8). The explanatory variable of main interest is the identity treatment interacted with Clicked and interacted with GermanLink ("IdentityxClickedxGermanLink") which takes the value of 1 if learner i participated in the identity treatment and clicked on the German Link in at least one of the emails and 0 otherwise. Days since registration measures the time a learner was registered on the platform before the first intervention email was sent. Active before intervention is a dummy variable that takes the value of 1 when a learner spent time on the platform before the intervention and 0 otherwise. The dummy variable Female takes the value of 1 when learner i is female and is 0 otherwise. Heteroscedasticity-robust Huber-White standard errors are in parentheses.

Table VII: Participant Survey Overview

	N	Mean	SD	Min	Max
<i>Socio-economic characteristics</i>					
Age	721	29.997	7.832	12	63
Partner (d)	782	0.474	0.500	0	1
Number of Children	778	0.814	1.300	0	9
Years of schooling	634	13.692	3.614	9	19
Graduated secondary school (d)	637	0.914	0.281	0	1
Want German Citizenship (d)	653	0.989	0.103	0	1
<i>Country of Origin</i>					
Syria (d)	786	0.714	0.452	0	1
Afghanistan (d)	786	0.050	0.217	0	1
Iran (d)	786	0.048	0.215	0	1
Iraq (d)	786	0.033	0.179	0	1
Other (d)	786	0.902	0.297	0	1
<i>Religion</i>					
Muslim (d)	779	0.656	0.475	0	1
Christian (d)	779	0.132	0.339	0	1
Other (d)	779	0.087	0.282	0	1
<i>Asylum Status</i>					
Rejected (d)	494	0.024	0.154	0	1
No Deportation (d)	494	0.032	0.177	0	1
Recognition (d)	494	0.302	0.459	0	1
Protection (d)	494	0.162	0.369	0	1
Refugee (d)	494	0.480	0.500	0	1
Asylum Requested (d)	642	0.893	0.310	0	1
<i>Occupation (multiple answers possible)</i>					
Working	656	0.200	0.400	0	1
Internship	656	0.072	0.258	0	1
School/University	656	0.270	0.444	0	1
Waiting	656	0.291	0.455	0	1
Integration Course	656	0.364	0.482	0	1
Apprenticeship	656	0.081	0.273	0	1
<i>Language Skills</i>					
German: read and write (very) well	642	0.416	0.493	0	1
German: speak (very) well	642	0.428	0.495	0	1
Total score language test (in %)	480	0.219	0.097	.033	.6
<i>Labor market outcomes</i>					
Wage (net)	177	868.347	796.559	0	6500
Working (d)	656	0.200	0.400	0	1
Weekly hours	204	27.647	13.962	5	65
Applications last month	233	3.981	3.528	0	11
Interviews last month	227	1.007	1.522	0	11

Note: Mailing sample.

Table VIII: Labor Market Outcomes

VARIABLES	(1) Working (d)	(2) Wage (net)	(3) Weekly hours	(4) Applications last month	(5) Interviews last month
Identity treatment (d)	-0.007 (0.033)	-29.905 (139.288)	-0.296 (1.976)	0.388 (0.495)	-0.142 (0.216)
Months since registration	0.005 (0.004)	11.516 (14.488)	0.239 (0.246)	0.010 (0.062)	0.034 (0.023)
Active before treatment (d)	0.070** (0.031)	144.410 (112.444)	-0.827 (2.211)	0.451 (0.491)	0.289 (0.192)
Female (d)	-0.054 (0.038)	-89.478 (154.915)	-2.657 (2.909)	-1.502** (0.581)	-0.268 (0.230)
Constant	0.104* (0.060)	641.624*** (217.722)	25.439*** (4.159)	3.485*** (0.939)	0.495 (0.318)
Observations	656	177	204	233	227
R-squared	0.013	0.013	0.009	0.026	0.021
RI p-value of ident. treat.	0.83	0.81	0.89	0.40	0.49
RI SE of p-value	0.0038	0.0039	0.0032	0.0049	0.0050
RI repetitions	10000	10000	10000	10000	10000

Note: Coefficient estimates from ordinary least squares estimations. * significant at 10%; ** sign. at 5%; *** sign. at 1%. All dependent variables are recored in a questionnaire that was send out 10 weeks after the intervention. The dependent variable reports if a learner states that she is employed in column (1). Other options were waiting for authorities, studying, being unemployed or doing nothing. For the subsample of the employed the dependent variable in column (2) reports the learner net wage, in column (3) how many hours the learner works. For the subsample of the unemployed the dependent variable in column (4) reports how many applications the learner send out in the past month and in column (5) how many interviews the learner attended. The explanatory variable of main interest is identity treatment which takes the value of 1 if learner i participated in the identity treatment and is 0 otherwise (control group - reminder only). Days since registration measures the time a learner was registered on the platform before the first intervention email was sent. Active before intervention is a dummy variable that takes the value of 1 when a learner spent time on the platform before the intervention and 0 otherwise. The dummy variable Female takes the value of 1 when learner i is female and is 0 otherwise. Heteroscedasticity-robust Huber-White standard errors are in parentheses. Additionally, randomization inference based p-values, their standard error and the number of permutation to compute these p-values are reported for the identity treatment dummy at the bottom of the table. This takes into account the triple-wise stratification of the intervention.