Lifting the Curtain: Backstage Cognition, Frontstage Behavior, and the Interpersonal Transmission of Culture

Richard Lu¹, Jennifer A. Chatman¹, Amir Goldberg², and Sameer B. Srivastava¹

¹Haas School of Business, University of California, Berkeley
²Stanford Graduate School of Business

October 23rd, 2017

Abstract

From the schoolyard to the boardroom, the pressures of cultural assimilation pervade all walks of social life. Yet people vary in the capacity to fit in culturally, and their fit can wax and wane over time. We examine how individual cognition and social influence produce variation and change in cultural fit. We do so by lifting the curtain between the backstage (cognition) and frontstage (behavior) of cultural fit. We theorize that the backstage comprises two analytically distinct dimensions—perceptual accuracy and value congruence—and that the former matters for normative compliance on the frontstage, whereas the latter does not. We further propose that a person’s behavior and perceptual accuracy are both influenced by observations of others’ behavior, whereas value congruence is less susceptible to peer influence. Drawing on email and survey data from a mid-sized technology firm, we use the tools of computational linguistics and machine learning to develop longitudinal measures of frontstage and backstage cultural fit. We also take advantage of a reorganization that produced quasi-exogenous shifts in employees’ peer groups to identify the causal impact of social influence.

Keywords—Culture, Person-Culture Fit, Cognition, Behavior, Language, Machine Learning, Organizations.

Introduction

Whether assimilating to a country or adapting to a new school, people typically seek to fit in culturally with their social groups. The benefits of conformity, as well as the sanctions and penalties that come with failed cultural integration, are particularly stark in contemporary organizations. Indeed, prior work has consistently demonstrated that high levels of individual cultural fit are associated with increased productivity, stronger
commitment, and less turnover (Kristof-Brown, Zimmerman, and Johnson 2005; Chatman and O’Reilly 2016). Moreover, employers have increasingly emphasized screening, selecting, and socializing new hires on the basis of cultural fit rather than focusing exclusively on hiring for skills (Chatman 1991; Rivera 2012; Meyer, Hecht, Gill, and Toplonytsky 2010). At the same time, as the average tenure in firms has declined (Hall 1996), workers must frequently retool themselves culturally as they move from one organization to the next. Yet people vary considerably in their ability to adapt culturally within a given organization (Srivastava, Goldberg, Manian, and Potts 2017). Why do some people fit in better than others and why do people vary in their ability to adjust their cultural fit over time?

Prior work has implicitly assumed that individuals who fit into their social environments both think and act in ways that are consistent with their peers’ thoughts and behavioral expectations. Yet, even when people feel pressure to fit in, how they think and feel about their social groups can often differ substantially from how they behave when interacting with other members. This distinction between backstage cognition and frontstage behavioral presentation—to borrow and adapt Goffman’s (1959) dramaturgical metaphor—is consequential because behavior is public but cognition is private. As a result, people can either purposefully or unintentionally act in ways that lead others to misperceive their underlying thoughts and intentions (Gardner and Martinko 1988; Hochschild 2012). To understand how cultural fit varies across people and changes over time, we contend that it is necessary to lift the curtain that typically separates the backstage from the frontstage.

Integrating insights from cultural sociology and organizational psychology, we uncover a heretofore obscured area of the backstage. Although prior work has thought of the cognitive dimension of fit as the match between individuals’ values and those of the social group to which they belong (Alba and Nee 2009; Chatman 1989), we propose a novel and complementary conceptualization of backstage cultural fit: the degree of

---

1We follow Goffman (1959: 107) in conceptualizing the frontstage as the behavior a person exhibits to “give the appearance that his activity...maintains and embodies certain standards.” In similar fashion, we view the backstage as the private thoughts a person has when he can “drop his front, forgo speaking his lines, and step out of character” (Goffman, 1959: 112). Whereas Goffman’s conception of the backstage includes both private thoughts and unobserved actions—for example, how staff in a mental hospital might feel about a difficult patient and talk to him outside of visiting hours when family members are not present—we instead focus on just the cognitive component of the backstage. Specifically, we use the term backstage to refer to people’s inner thoughts about the degree of cultural alignment between themselves and their social group.
correspondence between how individuals and their peers perceive their cultural milieu. In other words, we propose that backstage cultural fit encompasses two different underlying components: value congruence, or the alignment between what an individual values in an organization and what her peers value; and perceptual accuracy, or an individual’s ability to accurately decipher the cultural code of her social group.

We further propose that these two forms of backstage cultural fit have differing consequences for individual outcomes in the organization. Consistent with prior research (e.g., [Meglino and Ravlin, 1998]), we suggest that value congruence is a relatively stable aspect of cognition that does not directly influence one’s capacity to interact with peers in normatively compliant ways but instead predicts a person’s self-identification and long-term attachment to the organization (O’Reilly and Chatman, 1986). In contrast, we theorize that perceptual accuracy, a more pliable dimension of backstage cultural fit, is intimately tied to behavioral conformity on the frontstage. Specifically, we propose that increases or decreases in perceptual accuracy produce corresponding shifts in a person’s ability to act in normatively compliant ways. We further argue that these changes in perceptual accuracy arise, in part, through observations of others: witnessing normatively compliant behavior in peers boosts one’s own perceptual accuracy and hence one’s capacity for normative compliance.

To evaluate these ideas, we draw on survey data, eight years of internal email data, and personnel records from a mid-sized technology firm. We use the tools of computational linguistics and machine learning to transform the cross-sectional measures of backstage cultural fit, which were assessed through a survey instrument, into longitudinal measures and to develop measures of frontstage cultural fit based on the linguistic style that employees used in email communications with their colleagues. We also take advantage of a reorganization that produced quasi-exogenous shifts in employees’ peer groups to identify the impact of social influence—that is, of how a focal actor’s perceptual accuracy and behavioral fit changed in response to essentially random changes in the peers to which she was connected. Findings from this investigation shed new light on the relationship between cognition and behavior and the interpersonal transmission of culture.
Cultural Fit on the Backstage and Frontstage

Arguments about culture typically make implicit assumptions about underlying cognitive processes (DiMaggio, 1997). Sociologists often define culture as “shared understandings,” namely, similarities between individuals’ beliefs, value systems, and interpretations. In most everyday settings, one’s private cognition is, however, unavailable to others. Rather, one observes others’ behavior and then draws inferences—with varying degrees of accuracy—about their beliefs, values, and motivations (Kelley and Michela, 1980).

Culture, in other words, resides both in the distribution of inner thoughts and observable behaviors across individuals. Cultural fit, by extension, can be thought of as comprising two related but distinct dimensions: backstage (or cognitive) cultural fit, which relates to the degree of shared understanding between an individual and her peers, and frontstage (or behavioral) cultural fit, or the extent to which an individual’s behaviors are compliant with the group’s normative expectations (Mobasseri, Goldberg, and Srivastava, forthcoming).

Previous work has tended to focus on either cognitive or behavioral fit and implicitly assumed that the two correspond highly. An extensive literature in organizational psychology has, for example, examined culture through the lens of person-environment fit, highlighting the importance of shared values among organizational members (Edwards and Cable, 2009; Ostroff and Judge, 2007). This work has primarily illuminated the backstage of cultural fit. In doing so, it has identified two core mechanisms that link the backstage to individual attainment. The first relates to self-perceptions. Individuals whose values are

---

2The definition of culture as an analytical construct has long been a matter of debate by sociologists, and we do not attempt to fully resolve this debate here. “Shared understandings,” in our view, is a useful shorthand in that it points to two important properties of culture: that it dwells in the similarities between the individuals who constitute a group and that these similarities relate to group members’ mental representations of the world, often manifested in the group’s patterns of behavior. Missing from this useful, albeit simple, definition is the idea that such shared understandings emerge through interpersonal interaction.

3We acknowledge that not all individuals seek to fit in behaviorally and that some people are more predisposed than others to engaging in non-compliant behavior. Although the need for uniqueness is most likely hard-wired, it is also balanced by the propensity for compliance and assimilation with important social groups (Leonardelli, Pickett, and Brewer, 2010). Moreover, the tendency to conform is mediated by individual endowments: those with high status or who enjoy structural buffering by virtue of being embedded in a tight-knit community may under some circumstances reap the benefits of culturally non-compliant behavior while limiting its adverse consequences (Goldberg, Srivastava, Manian, Monroe, and Potts, 2016). On balance, however, behavioral conformity is generally beneficial such that people are by and large motivated to conform to the normative expectations of their social group (Miller and Prentice, 2016). Thus, we expect individuals to be attuned to their cultural environments and to respond to their peers’ behaviors in their attempts to fit in.
compatibility with those prevalent in an organization are more likely to self-identify with that organization (Cable and Judge, 1996; Judge and Cable, 1997; O’Reilly and Chatman, 1986). Such identification, in turn, leads to greater attachment, heightened motivation, stronger commitment, and higher productivity (Chatman, 1991). The second relates to the ease of interpersonal interaction and coordination. Culturally aligned individuals find it easier to interact with one another because they have mutually compatible expectations of behavior (Ellenbein and O’Reilly, 2007; Morrison, 2002).

Yet, in many cases, people can successfully interact with one another even when they do not share the same values. Work in organizational psychology (Hewlin, 2003; Hewlin, Dumas, and Burnett, 2017) and sociology (Hochschild, 2012) finds that people often behave in ways that are consistent with their social group’s normative expectations even when these norms are incompatible with their own private beliefs. As Willer and his colleagues (2009) demonstrate, this ability to separate beliefs from behaviors can lead to the persistence of unpopular norms. The core distinction is between the beliefs people value personally and those they perceive to be widespread in the social group (cf. Goldstein, Cialdini, and Griskevicius, 2008). When group members believe that a behavior is prevalent—and consequently falsely infer that associated privately held values are also widespread—they often sanction those who fail to conform. The fear of being exposed as inauthentic or deviant motivates them to police the cultural order despite their private disagreement with it.

To understand how such a situation can arise, it is important to distinguish between two dimensions of cognition: preferences and construals. Whereas preferences define which behaviors are desirable, construals refer to the levels of abstraction and the associated mental representations that a person conjures when making sense of a situation. How an individual construes a social setting affects which of her preferences will be activated and ultimately what action she will pursue (Trope and Liberman, 2010). Shared understandings do not necessarily require that all group members hold the same preferences. Rather, to share understandings is, first and foremost, to construe daily experiences through similar interpretative lenses (Goldberg, 2011; DiMaggio and Goldberg, forthcoming).

Goffman (1959) makes a comparable point in his work on interpersonal interaction. As long as group members have a shared understanding of a situation—of the social roles it implies, of the behaviors
appropriate to those roles, and of the implicit meanings these behaviors convey—then interactions between
them can occur rather seamlessly. Further, even when there is agreement on how a situation is construed,
group members can still craft their self-presentations in a manner that decouples their behavior from their
privately held preferences. In the absence of such agreement, however, interaction breaks down, leading
to incompatibilities between one person’s expectations and another’s behavior. Under such circumstances
backstage cognition is more likely to leak unintentionally into frontstage behavior.

Two Areas of the Backstage: Value Congruence and Perceptual Accuracy

Preferences and construals are aspects of individual cognition; however, they become culturally
meaningful when we consider the individual in relation to her social group. Value congruence represents
the cultural manifestation of preferences in that it reflects the match between what individuals prefer and
what prevails in the social group. Perceptual accuracy is instead the cultural analogue of construals in that it
indicates the degree of alignment between a person’s perceptions and those of other group members.

More specifically, value congruence is the degree of similarity between an individual’s own preferred
values and those reported by others as being prevalent in the group. By “value,” we mean enduring beliefs
about desired or undesired ways of working and interacting with others (e.g., “I prefer a friendly work
environment”), as distinguished from situation-specific preferences (e.g., “I prefer having lunch before noon”) (O’Reilly, Chatman, and Caldwell, 1991). Note that value congruence relates to fit with the normative
environment, irrespective of whether other group members privately hold the same preferences. In an
“Emperor’s New Clothes” dynamic of the kind that Centola, Willer, and Macy (2005) discuss, a person might
have low value congruence if she prefers not to blindly defer to hierarchy when the prevalent behavioral norm
is to defer to more senior others.

People whose ideal preferences are compatible with those prevalent in their social environment find
it easier to maintain a positive self-concept (Chatman and Barsade, 1995). Consequently, they identify
more strongly with the organization and derive greater satisfaction from their interactions with others. We
therefore expect value congruence to be primarily related to motivation and long-term attachment to the
Lifting the Curtain

organization—as evidenced by a negative association between value congruence and the choice to exit the organization voluntarily.

We do not, however, anticipate that value congruence will be consequential for a person’s capacity to conform to her group’s normative expectations of behavior. Although people whose values are more congruent with their organization’s may be motivated to behave in normatively compliant ways, they may still lack the information needed to do so. It is one thing to prefer, for example, a cooperative work environment and another to understand which behaviors signal cooperativeness in a specific normative context.

As a complement to value congruence, we introduce a distinct component of backstage cultural fit: *perceptual accuracy*, or the degree to which an individual correctly reads the group’s cultural code. We define perceptual accuracy as the extent to which an individual’s assessment of the behaviors that are or are not normatively compliant with group members’ expectations is consistent with the readings of her peers. Note that this accuracy does not relate to peers’ private beliefs or preferences. In the “Emperor’s New Clothes” situation, a perceptually accurate individual will correctly observe that the appropriate behavior is to express admiration for the monarch’s clothes, irrespective of whether she correctly perceives that the majority of her peers believe that the emperor is, in fact, naked. As we detail below, we conceptualize and operationalize perceptual accuracy at a high level of construal that relates to group norms and, as such, usefully informs members’ behaviors across many relevant group situations.

**Perceptual Accuracy and Frontstage Behavioral Fit**

Perceptual accuracy relates to an individual’s ability to decipher the cultural code implicit in others’ behaviors. Although organizations often formalize their idealized values into cultural statements, interpreting the local normative environment is a subtle, complex, and ongoing cognitive task. A colleague’s cynical joke in a meeting, for example, can be interpreted as a friendly attempt to establish rapport or as a derogatory comment aimed at undercutting others. Correctly construing this behavior requires tacit and layered knowledge that connects behaviors, symbols, and meanings to abstract cultural categories. Possessing this knowledge is essential to knowing how to behave appropriately. Perceptual accuracy is, we argue, intimately related to the
capacity to behave in culturally compliant ways with one’s peers.

Figure 1 illustrates these conceptual arguments and their behavioral implications. Imagine five possible values (labeled a to e) that people can espouse. The four individuals depicted in the diagram (labeled A to D) correspond to four hypothetical organizational members. Each individual is characterized by three distributions: her (1) values (V) and (2) perceptions (P) on the backstage; and her (3) behaviors (B) on the frontstage. As noted above, only the behaviors of others are directly observable; their values and perceptions are private and can only be indirectly inferred.

[FIGURE 1 ABOUT HERE]

Individual A in Figure 1 is perceptually accurate but value incongruent: her perceptions of the cultural code (P) are consistent with the majority of her peers’, but the prevailing values are mostly inconsistent with her own (V). Nevertheless, her behavior mirrors her perceptions. Suppose that value d is confrontation-orientation. Although A does not prefer a confrontational environment, she sees confrontation as a common and legitimate behavior in the organization. She is consequently likely to express disagreement and negation in her interaction with others (as reflected in her tendency to exhibit behavior d). Individual D, in contrast, is also conflict-averse, but unlike A she misperceives the prevalence of confrontation in the organization. Consequently, her behavior is incongruent with her peers’. She is more likely to be accommodating and apologetic, whereas her peers are confrontational. Although the four hypothetical individuals in the diagram espouse different values, only D is a behavioral misfit. Like A, individuals B and C behave in a normatively compliant way because they hold similarly accurate perceptions of the cultural code despite the latter two being more value congruent than the former.

In sum, we argue that one dimension of backstage cultural fit—perceptual accuracy—is closely linked to an individual’s capacity for frontstage cultural fit, whereas the other dimension—value congruence—is related to self-identification and long-term attachment but not to contemporaneous behavioral cultural fit. We therefore hypothesize:

Hypothesis 1: Perceptual accuracy is positively related to behavioral cultural fit.
The Interpersonal Transmission of Culture

Contending that perceptual accuracy, rather than value congruence, predicts behavioral cultural fit shifts the analytical focus from heterogeneity between individuals’ preferences and beliefs to differences in their ability to enculturate—that is, their ability to read and adapt to the cultural code. A dominant line of work on culture in organizations conceptualized cultural fit as a fundamental compatibility between individuals and organizations—a match between the “personalities” of the individual and the group (Cable and Judge, 1996; Schneider, 1987). This perspective continues to implicitly guide personnel practices in the contemporary workplace. Many organizations emphasize cultural fit in the hiring phase, assuming that only certain individuals possess innate qualities or underlying values that make them a strong cultural match (Rivera, 2012). Yet cultural fit is a dynamic process: individuals are capable of adapting their behavior to the prevailing norms in an organization (Van Maanen and Schein, 1979; Chatman, 1991; Srivastava et al., 2017). People acquire this capability through ongoing socialization into the organization (Ashforth and Saks, 1996; Van Maanen, 1975).

What factors lead some people to increase their level of behavioral cultural fit over time, while others remain stagnant? One line of work attributes such variance to psychological differences between individuals. For example, a robust literature in social psychology has focused on self-monitoring orientation—a sensitivity and responsiveness to social cues of situational appropriateness (Kilduff and Day, 1994; Snyder, 1979). High self-monitors tend to regulate their behavior given their read of what is expected of them, whereas low self-monitors hew to their sense of self, irrespective of the situation. Self-monitoring is also related to a capacity for deep-acting, the ability to adapt emotions to organizational expectations, leading to more genuine displays of cultural congruence (Grandey, 2000; Scott, Barnes, and Wagner, 2012). High self-monitors, in other words, are more motivated to read the cultural code, more inclined to conform to it, and more likely to be perceived as authentic when they do.

Yet perceptual accuracy is also a matter of situational context, not just of intrinsic ability. By virtue of
their innate “cultural intelligence” (Tomasello, 2009), human beings are intrinsically motivated to be attuned to the cultural code prevalent in their social environments. In other words, we expect perceptual accuracy to be a pliable dimension of cognitive cultural fit. Adjusting to the cultural code of a group is, by definition, a process of social learning. The quality of this learning depends not only on the student, but also on the peers from whom she learns.

We therefore expect that the composition of a person’s network has a bearing on her ability to correctly read the cultural code and to adapt her behaviors accordingly. Experimental work in young children, for example, demonstrates that exposure to multiple and consistent behaviors increases the fidelity and speed of cultural transmission (Henrich, 2016; Herrmann, Legare, Harris, and Whitehouse, 2013). Similarly, in the workplace, employees’ ability to learn and their susceptibility to influence from others is related to the kinds of colleagues with whom they interact (Chan, Li, and Pierce, 2014; Liu and Srivastava, 2015). In particular, having colleagues who themselves have a more accurate read of the cultural environment can help correct one’s own misperceptions, thereby improving one’s own perceptual accuracy (Balkundi and Kilduff, 2006).

Importantly, people primarily have access to their peers’ fronstage behavior. It is through observing these behaviors that they develop their own perceptions of the cultural environment. We therefore anticipate that peers’ fronstage behavior—as opposed to their backstage cognition—will influence the focal individual’s own thoughts and behavior. Moreover, because we argue that the ability to behave compliantly is primarily dependent on perceptual accuracy, we also expect that individuals’ perceptual accuracy will be influenced through their observations of their colleagues. In contrast, we argued above that value congruence is not directly linked to frontend behavior. It is also likely, we propose, to remain relatively stable given that individuals’ deeply held values are encoded in implicit cognition and thus slow to change (Meglino and Ravlin, 1998; Vaisey, 2009; Srivastava and Banaji, 2011). We therefore expect that value congruence will be less susceptible to peer influence than will perceptual accuracy.

In support of these expectations, an extensive literature has shown that individuals’ attitudes can change as a direct consequence of exposure to and interaction with their network contacts (Baldassarri and Bearman, 2007; Friedkin and Johnsen, 1990; Marsden and Friedkin, 1993); however, exposure to peers whose deeply
held values and beliefs run counter to one’s own can also activate biases in information processing such that discordant information is discounted or even rejected (Lord, Ross, and Lepper, 1979; Dandekar, Goel, and Lee, 2013). In contrast, expectations of normatively appropriate behavior are strongly shaped by shared perceptions that arise through interaction and observation (Friedkin, 2001). Taken together, these findings lead to the prediction that a person’s perceptions of the cultural order will be more susceptible to social influence than will her deeply rooted values, beliefs, and preferences.

The causal assumptions informing this model are depicted in the arrows in Figure 1. Individual A observes B’s behavior and updates her perceptions accordingly. These perceptions, in turn, affect how she behaves. Her values, in contrast, remain relatively unchanged. Overall, we expect:

**Hypothesis 2:** Perceptual accuracy and behavioral fit are both susceptible to behavioral peer influence. Specifically, as one’s peers behave in more (less) normatively compliant ways, one’s own perceptual accuracy increases (decreases) and one’s behavioral fit concomitantly increases (decreases).

**Methods**

Testing these hypotheses requires access to longitudinal data on backstage and frontstage cultural fit, as well as exogenous variation in the set of peers to which a focal actor is exposed. Prior work has derived point-in-time snapshots of backstage cultural fit based on responses to culture surveys (Chatman, 1991) and, more recently, developed ways of measuring time-varying frontstage cultural fit based on the content of communication among colleagues (Goldberg et al., 2016; Srivastava et al., 2017). As described in more detail below, we conducted a cultural survey in a mid-sized organization and analyzed the email communications of its members. We used the tools of computational linguistics and machine learning to impute time-varying measures of perceptual accuracy, value congruence, and behavioral fit for all employees, including those who chose not to participate in the culture survey, and for all time periods for which we have email data. We use both cross-sectional data and these imputed longitudinal measures to test the first hypothesis. In addition, we
exploit a reorganization event that produced quasi-exogenous shifts in employee’s peer groups to derive a causal test of the second hypothesis.

Data

Our empirical setting is a mid-sized technology firm. We obtained three types of data: (1) personnel records provided by the firm, (2) email communications between organizational members, and (3) an indirect self-report, the Organizational Culture Profile (OCP), which employees of the firm completed (Chatman, Caldwell, O’Reilly, and Doerr 2014). Once we matched the raw data and removed identifying information, the resulting data set consisted of 29,255 person-month observations, spanning the period from 2008 to 2016.

**Personnel Records.**—We obtained monthly extracts from the firm’s human resource information system. These extracts included demographic information such as age and gender, organizational status such as departmental affiliation and start date, and information about individual outcomes such as monthly bonus received, date of exit, and reason for exit (voluntary or involuntary).

**Email Data.**—We collected eight years of email data from the organization, including not only metadata (i.e., who sent messages to whom and when) but also raw message content. Given our focus on cultural dynamics within the organization, we excluded emails exchanged between employees and the outside world. We also eliminated automatically generated messages and, per instructions from the company’s in-house lawyers, messages sent from or to members of the (small) legal department. The resulting data set included over five million unique emails.

**Organizational Culture Profile.**—To assess backstage cultural fit, we implemented the widely used Organizational Culture Profile (Chatman et al. 2014). The OCP provides a quantitative assessment of an organization’s culture. The OCP, which has been extensively used in organizational research, consists of 54 norm statements (e.g., fast moving, being precise) that emerged from a review of academic and practitioner-oriented writings on culture (O’Reilly et al. 1991, Sarros, Gray, Densten, and Cooper 2005). We sent two versions of the OCP to the organization, one asking employees to characterize the current culture of the organization and the other asking employees to characterize their personally desired culture. All
employees completed the survey describing the organization’s current culture and a randomly selected half of employees completed the survey of their own personally desired cultural characteristics. Overall, we received 440 completed surveys about the current organizational culture and 238 completed surveys about the personally desired culture.

Variables

The Frontstage: Behavioral Cultural Fit

We operationalized behavioral fit using the Interactional Language Use Model (Goldberg et al., 2016; Srivastava et al., 2017), which assesses the extent to which an individual’s style of communication matches that of her interlocutors. Employees exhibiting high behavioral fit communicate in ways that align with their colleagues’ communication styles—for example, in the use of personal pronouns or emotional display (e.g., expressions of positive or negative affect) (Doyle, Goldberg, Srivastava, and Frank, 2017; Bail, Brown, and Mann, 2017).

To derive this measure, we first translated raw email content into a format that captures communication style, as reflected in the well-established and widely used Linguistic Inquiry and Word Count (LIWC) lexicon (Pennebaker, Booth, and Francis, 2007). LIWC is a semantic dictionary that maps words to 64 distinct emotional, cognitive, and structural categories. To transform raw emails to a measure of linguistic behavioral fit, we followed the procedure outlined by Goldberg et al. (2016) and Srivastava et al. (2017).

Specifically, we define $\vec{m}_{it}$ as each email individual $i$ sends at time $t$ and $\overleftarrow{m}_{it}$ as each email individual $i$ receives at time $t$. We then define the set of LIWC categories as $L$ and the set of all times in any given month as $T$. Our procedure iterates over all emails sent and received and produces $\vec{m}_{it}^l$ and $\overleftarrow{m}_{it}^l$ for the count of terms in email $\vec{m}_{it}$ and $\overleftarrow{m}_{it}$ in LIWC category $l \in L$, respectively. Then, by aggregating all individual email counts $\vec{m}_{it}^l$ and $\overleftarrow{m}_{it}^l$ for $t \in T$, it produces sent and received LIWC counts in month $T$, $\vec{m}_{iT}^l$ and $\overleftarrow{m}_{iT}^l$. We

---

4The other half completed a survey of the cultural characteristics needed for the organization to be successful in the future. We shared the results of this latter survey with organizational leaders as a condition of gaining access to the organization as a research site; however, we do not report these results here because they do not pertain to our theory and hypotheses.

5We acknowledge that there are many possible manifestations of behavioral conformity, including styles of dress and nonverbal communication. We focus on one important facet of normative compliance—linguistic fit—which we can observe over time for virtually all individuals in the organization.
normalize each LIWC count in each month by the total of all LIWC counts in that month to transform the LIWC probability distribution to a standard probability distribution. We use the notation, $O^i_{IT}$ to denote the outgoing normalized probability and $I^i_{IT}$ to denote the incoming normalized probability.

$$O^i_{IT} = \frac{\bar{m}^i_{IT}}{\sum_{l \in L} \bar{m}^i_{IT}}$$

(1)

$$I^i_{IT} = \frac{\bar{m}^i_{IT}}{\sum_{l \in L} \bar{m}^i_{IT}}$$

(2)

We define an individual $i$’s behavioral fit in month $T$ as the negative log of the Jensen-Shannon (JS) divergence (Lin 1991) metric between $i$’s outgoing and incoming normalized distributions:

$$BF_{IT} = -\log (JS(O^i_{IT} \parallel I^i_{IT}))$$

(3)

where the JS-divergence between two probability distributions is defined as a symmetric measure built by first taking the mean probability distribution between the normalized outgoing and incoming distributions, $M_{IT} = \frac{1}{2}(O_{IT} + I_{IT})$, and summing the Kullback-Leibler (KL) divergence (Kullback and Leibler 1951) of the outgoing and incoming distributions from that mean probability distribution.

$$JS(O^i_{IT} \parallel I^i_{IT}) = \frac{1}{2}KL(O^i_{IT} \parallel M_{IT}) + \frac{1}{2}KL(I^i_{IT} \parallel M_{IT})$$

(4)

$$KL(D^i_{IT} \parallel M_{IT}) = \sum_{l \in L} D^i_{IT} \log_2 \frac{D^i_{IT}}{M^i_{IT}}$$

(5)

Intuitively, when the outgoing and incoming distributions are nearly identical, the divergence approaches zero, suggesting a high level of behavioral fit; conversely, greater deviation between the probabilities of usage of LIWC categories translates to greater divergence and thus implies lower levels of behavioral fit. Stated differently, the more an employee’s use of cognitive, emotional, and structural terms in sent emails matches
Lifting the Curtain

the use of those terms in received emails, the greater her behavioral fit in a given month. Note that this
measure of fit assumes that the salient reference group for a focal individual is the set of colleagues with
whom she has email contact in a given month. Given that subcultures in organizations do not necessarily
conform to the contours of formal subunits, this choice of reference group allows us to identify a person’s fit
in an empirically grounded manner, without having to make assumptions about the boundaries of subcultures
in the organization.

The Backstage: Perceptual Accuracy and Value Congruence

We operationalized perceptual accuracy and value congruence based on employee responses to an
indirect self-report: the Organizational Culture Profile (OCP) (Chatman et al., 2014). The OCP is based on the
Q-sort methodology (Block, 1961). Respondents were asked to rank 54 value statements into nine categories,
with a specified number of statements in each category. This sorting of value statements represents a cultural
profile. To derive measures of fit from these profiles, we then calculated the correlation between profiles by
translating each value statement into its corresponding category number. For example, if value statement 1
were put in category 7 in one profile and category 2 in another profile, that statement would represent point
(7,2). We similarly computed points for all 54 value statements and used the correlation among those points.

We configured the OCP to yield two separate culture profiles for each respondent: a profile based on
her assessment of the current organizational culture and a profile based on her preferences for each value
statement. For the former, we asked: “To what extent do the value statements characterize the organization as
a whole?” For the latter, we asked: “To what extent do the value statements characterize your personally
desired values, that is, the values you desire in an organization?” Our two measures of backstage cultural fit
are based on the correlation between individual \(i\)’s cultural profile and a reference group cultural profile. To
make these measures comparable to our measure of behavioral fit, we chose the same reference group—i.e.,
the set of colleagues a person had email contact with in a given month weighted by communication volume.
We defined \(\text{perceptual accuracy}\) as the congruence between an individual’s current culture profile and the

---

\(^6\)The required distribution of statements across categories that range from least to most characteristic of a given value is 2-4-6-9-12-9-6-4-2.
reference group’s current culture profile. Similarly, we defined \textit{value congruence} as the correspondence between an individual’s personal culture profile and the reference group’s current culture profile. Note that the reference group profile is identical in both cases. The difference between the two measures stems from the choice of individual culture profile: current culture for perceptual accuracy and personal culture for value congruence. For robustness checks reported below, we also produced versions of these measures in which the reference group included all employees in the organization rather than just the focal individual’s email interaction partners in a given month.

\textit{Imputing Backstage Cultural Fit Over Time}.—The procedure above creates cross-sectional measures of perceptual accuracy and value congruence; however, to convincingly test hypotheses about the dynamic interrelationships among the three fit measures, we need longitudinal backstage measures. Taking inspiration from Salganik’s (2017) notion of \textit{amplified asking}—that is, combining surveys with digital trace data to infer responses for people who cannot be feasibly surveyed or whose responses are missing—we undertook a procedure based on machine learning techniques to identify from raw email content (rather than the higher-level LIWC categories used to derive our measure of behavioral fit) the “linguistic signature” of perceptual accuracy and value congruence (see also Bail, 2017).

We assumed that, if language reflects internal processes of cognition (Pinker, 2007), then there should be an identifiable relationship between email communication and backstage cultural fit. If this relationship can be discerned through machine learning, then it should be possible to impute perceptual accuracy and value congruence measures for all employees, including those who departed before the OCP was implemented and those who were employed but chose not to participate. Moreover, assuming a relatively stable underlying relationship between language use and cognition, these measures can be imputed for individuals at all points in time for which they exchanged email messages with colleagues. In other words, this procedure allowed us to transform a one-time collection of value preferences and perceptions of the current culture, based on the OCP, into longitudinal measures of backstage cultural fit.

We used a random forest model to help uncover this underlying link between language and cognition (Ho, 1995; Friedman, Hastie, and Tibshirani, 2001). Random forest models have several beneficial characteristics
for this task: they can detect arbitrary, nonlinear relationships; they typically require fewer observations than do other machine learning methods to produce comparable results; and they are inherently robust to overfitting, or incorrectly inferring signal from idiosyncratic noise in the data. Figure 2 provides a conceptual overview of this procedure. Further procedural details are provided in the Appendix A; evaluative analyses regarding model fit are provided in Appendix B.

Peer Cultural Fit

After imputing perceptual accuracy and value congruence, we turned next to identifying the distribution of these measures in the network of email contacts surrounding a focal individual. To do this, we first identified an individual $i$’s communication partners $J$ for each month $T$. Then, using our time-varying measures of backstage cultural fit, as well as our time-varying measure of behavioral fit, we took the mean cultural fit for all communication partners $J$, weighted by the volume of incoming communication received from each interlocutor, to generate $i$’s peer cultural fit for month $T$. We did this for each cultural fit measure, yielding network-based measures that we refer to as peer behavioral fit, peer perceptual accuracy, and peer value congruence.

Individual Outcomes

To establish the validity of our imputed longitudinal measures, we implemented supplemental analyses reported below. These were not direct tests of our hypotheses but designed to assess whether the imputed measures related to career outcomes as would be expected based on theory and prior research. In particular, we derived from the personnel records two individual outcome measures. The first was monthly bonus. Only those in job roles such as sales or operations, for which productivity could be objectively assessed, were bonus eligible. For each of these roles, the company established a formula that linked specific productivity indicators—for example, a sales person’s conversion of leads into revenue—to monthly bonus payments. Given that the distribution of bonuses was skewed, we logged this measure in the analyses reported below. The...
second outcome was exit, based on an employee’s departure date. We used company records to distinguish between voluntary and involuntary exit.

Control Variables

We estimated both within-person and between-person models for our analyses. In within-person models, because time-invariant effects are subsumed by individual fixed effects, we only included as control variables managerial status (lagged), tenure, and departmental affiliation. For our between-person models, we included additional control variables for age and gender.

Analytical Approach

We tested Hypothesis 1 using OLS regressions based on cross-sectional data, as well as fixed effect regressions based on longitudinal data, including the imputed measures of perceptual accuracy and value congruence. We standardized all variables in the regression models reported below.

To test Hypothesis 2, we identified the effect of changes in peer composition on the focal individual’s cultural fit measures—behavioral fit, perceptual accuracy, and value congruence. We began by estimating the following basic OLS model, with individual, department and year fixed effects:

\[
CF_{idt} = \beta_0 + \beta_1 \langle PeerCF \rangle_{idt-1} + \beta_2 |Peer|_{idt-1} + \eta X_{idt-1} + \beta_3 Year_t + \beta_4 Dept_d + \beta_5 Ind_i + \epsilon_{idt}
\]

where \(CF_{idt}\) is the relevant cultural fit measure (behavioral fit, perceptual accuracy or value congruence) for individual \(i\) in department \(d\) at time \(t\), \(\langle PeerCF \rangle_{idt-1}\) is the mean peer cultural fit at time \(t - 1\) weighted by number of incoming messages, \(|Peer|_{idt-1}\) is the number of peers at time \(t - 1\), and \(X\) are time-varying individual attributes. The inclusion of individual fixed effects accounts for stable variation between individuals, such as differences in innate psychological traits, experience, and preferences. Department and year fixed effects account, respectively, for differences between departments (e.g., different demographic compositions)

Managerial status and departmental affiliation can be estimated in fixed effect models because some employees get promoted from individual contributor to managerial roles and because some employees move across departments.

18
and periods (e.g. variation in turnover rates) that might systematically affect cultural fit.

We lag mean peer cultural fit and number of peers to ensure appropriate temporal ordering. Yet even with individual fixed effects and lagged predictors, this modeling approach does not yield causal estimates. It could be the case, for example, that individuals with high cultural fit seek to interact with equally culturally integrated individuals. In other words, this modeling approach cannot separate the effects of homophily from those that arise through peer influence.

To address this problem, we exploited a reorganization event that transpired over a period of two months, roughly seven years after the firm’s founding. Ideally, we would have identified an exogenous shock that assigned certain individuals to interact with a random set of new peers, while others retained their previous network contacts. Such a natural experiment would allow for causal identification of peers’ cultural fit on that of the focal individual. Although this reorganization was not entirely random, it was driven primarily by functional needs arising from rapid growth at the time, and it affected all employees to some extent. Moreover, unlike network changes generated by downsizing, it did not disproportionately affect low-performing or otherwise systematically similar peers. As such, the reorganization can be thought of as quasi-exogenous in that it introduced significant random variation in employees’ network compositions. Recognizing, however, that this event was not a pure natural experiment, we used an extension of an instrumental variable peer effects model first introduced by Waldinger (2012). Using a two-stage least-squares model, we first estimated the random variation in mean peer cultural fit and number of peers introduced by the reorganization, and we then used this estimate to predict subsequent changes in cultural fit.

In typical instrumental variable designs, the instrument is assumed to only affect the endogenous variable. In the present case, however, the reorganization also affected the focal individuals’ peers’ network compositions. Thus, peers also experienced shifts in their cultural fit, driven by changes in their own peer group after the reorganization and social influence from peers in the month of reorganization. To address this complexity, we follow Waldinger (2012) and use induced change in peer cultural fit, \( \tilde{\Delta}_t \langle \text{PeerCF} \rangle \), as an instrument. \( \tilde{\Delta}_t \langle \text{PeerCF} \rangle \) is the change induced by the reorganization between periods \( t - 1 \) and \( t \), assuming peer cultural fit had remained fixed at its pre-reorganization level. Defining the measure in this way allowed
us to account for the change in peer exposure stemming from the reorganization, while separating out its downstream effects on peers’ cultural fit.

In addition to induced change in mean peer cultural fit, we also measured the magnitude of change in network composition as an instrument. Let \( I_{it} \) be a vector of length \( N \) (total number of employees) wherein each cell \( I_{it}(j) \) corresponds to the number of messages that \( i \) received from interlocutor \( j \) during month \( t \). We define \( i \)'s network change at time \( t \) as the cosine distance between \( i \)'s vectors of incoming messages in two consecutive months:

\[
NC(I_{it}, I_{it-1}) = \cos(I_{it}, I_{it-1})
\]

where the cosine distance between two vectors \( p \) and \( q \) is defined as:

\[
\cos(p, q) = 1 - \frac{\sum_{j=1}^{N} p(j)q(j)}{\sqrt{\sum_{j=1}^{N} p(j)^2} \sqrt{\sum_{j=1}^{N} q(j)^2}}
\]

Because the number of messages is non-negative, this measure is bounded by 0 and 1.

We used these instruments—network change, induced change in mean peer cultural fit, and the interaction between the two—to estimate the model’s two endogenous variables, mean peer cultural fit and number of peers. In the first stage we estimated the following regressions:

\[
\langle PeerCF \rangle'_{idt} = \beta_0 + \beta_1 NC(I_{it}, I_{it-1}) + \beta_2 \tilde{\Delta} \langle PeerCF \rangle_{idt-1} + \beta_3 NC(I_{it}, I_{it-1}) \cdot \tilde{\Delta} \langle PeerCF \rangle_{idt-1} + \beta_4 Ind_{-i} + \epsilon_{it} \quad \text{(9)}
\]

\[
|\text{Peer}|'_{idt} = \beta_0 + \beta_1 NC(I_{it}, I_{it-1}) + \beta_2 \tilde{\Delta} \langle PeerCF \rangle_{idt-1} + \beta_3 NC(I_{it}, I_{it-1}) \cdot \tilde{\Delta} \langle PeerCF \rangle_{idt-1} + \beta_4 Ind_{-i} + \epsilon_{it} \quad \text{(10)}
\]

In the second stage we estimated cultural fit at time \( t + 1 \) (a month after the reorganization) with instrumented mean peer cultural fit and number of peers as independent variables. These models included
individual, department, and year fixed effects. We specified the second stage regression as:

\[
CF_{idt+1} = \beta_0 + \beta_1 \langle PeerCF \rangle_{idt} + \beta_2 \langle Peer \rangle_{idt} + \beta_3 Year_t + \beta_4 Dept_{id} + \beta_5 Ind_{i} + \eta X_{it} + \epsilon_{idt}
\]  

(11)

where \( X_{it} \) represents time-varying individual controls. We report results from eq. (11) in the tables below.

**Preliminary Analyses—Evaluating the Variables of Interest**

Before turning to our main results, we summarize three preliminary analyses that sought to evaluate the validity of the backstage and frontstage cultural fit measures, particularly the backstage measures that were imputed using the procedure described in Appendix A. First, given that we theorized that value congruence is relatively stable over time, whereas perceptual accuracy is more susceptible to change, we traced the two imputed measures over a person’s tenure in the organization. We restricted this analysis to the first 36 months of employment given that only about 10% of employees had tenure exceeding 36 months during our observation period. We separately estimated OLS and fixed effect regressions of the two backstage variables using indicators for each month (up to month 36 of employment). These results are depicted in Figure 3. According to both models, when employees first enter the organization, they have relatively high value congruence and relatively low perceptual accuracy. Through approximately the first year of employment, however, perceptual accuracy increases sharply and continues a more gradual ascent thereafter. In contrast, value congruence increases—albeit not as steeply—in the first four months of employment and then remains mostly stable over the remaining months. These results support our contention that value congruence is relatively stable, while perceptual accuracy is more malleable.

[FIGURE 3 ABOUT HERE]

Second, in Table I we report the results of OLS regressions with individual, department and year fixed effects, where the dependent variable is bonus (logged), and frontstage and backstage cultural fit measures which are used as independent variables—behavioral fit, perceptual accuracy and value congruence—are
lagged. The fixed effects specification with lagged predictors allows us to estimate the effects of within-person change in cultural fit on subsequent productivity.

Whether modeled independently or together, all three cultural fit measures are significantly positively related to productivity. Thus we find, consistent with prior work (Chatman, 1991; Srivastava et al., 2017), that alignment on the frontstage, as well as on the backstage, is positively related to positive job performance—even when we use imputed longitudinal measures of backstage fit. The coefficients for behavioral fit and perceptual accuracy are of similar magnitude. The two variables retain their significance even when included together in Model 4.

In contrast, the effect of value congruence on bonus is more modest. This result is consistent with our expectation that value congruence remains more stable over time. Given that the unwavering component of value congruence is subsumed in our individual fixed effect, it is not surprising that its time-varying component accounts for less of the variance in job performance.

Finally, in Table 2 we modeled voluntary exit from the organization as a function of value congruence and perceptual accuracy. Although people leave organizations for a variety of reasons, voluntary exit is most likely to be associated with declining attachment. The competing risks model reported in Table 2 is a survival model that extends the Cox Proportional Hazards model to the case of multiple failures. In our case, involuntary exit is the competing risk.

As Table 2 indicates, value congruence is associated with a decreased risk of voluntary exit, while perceptual accuracy is not. The importance of value congruence in affecting voluntary departures, based on the imputed longitudinal measure, is consistent with prior work based on a cross-sectional measure of value congruence that predicted departure from firms up to two years later (Chatman, 1991). Overall, these supplemental analyses help to validate the longitudinal fit measures derived from our imputation methodology.

---

8Because including period fixed effects produces unstable estimates in such a model, we instead include the number of employees in the organization as a control. This accounts for time-varying fluctuations in average value congruence due to firm growth or decline. To account for variation in the number of observations per individual (some individuals remain only a handful of months in the organization, whereas others stay for years) we use overall tenure as a sampling weight.

9Neither perceptual accuracy nor value congruence is significant in predicting involuntary exit when we use the same framework with voluntary exit as the competing risk.
Main Results

Table 3 provides a test of our first hypothesis: that perceptual accuracy predicts changes in behavioral fit, whereas value congruence does not. The dependent variable in all models is behavioral fit. The first three models report results from cross-sectional data where backstage fit measures—perceptual accuracy and value congruence—are derived directly from the Organizational Culture Profile (OCP). These two measures are imputed in the three longitudinal models that follow.

Models 1 to 3 report results from cross-sectional data, with behavioral fit averaged over the three months preceding the administration of the OCP. In support of Hypothesis 1, perceptual accuracy is significantly related to behavioral fit, while value congruence is not; moreover, these patterns hold whether the two predictors are modeled separately (Models 1 and 2) or together (Model 3).

Table 3, Models 4 to 6, echo the results from the cross-sectional analyses in longitudinal specifications that include individual, department, and year fixed effects. Whereas the cross-sectional models estimate between-person effects, the longitudinal models with individual fixed effects estimate the relationship between within-person changes in backstage fit—perceptual accuracy or value congruence—and changes in behavioral fit. The longitudinal results provide further support for Hypothesis 1 given that perceptual accuracy is significantly related to behavioral fit, while value congruence is not. As individuals' perceptual accuracy increases, their behavioral fit correspondingly increases. Changes in value congruence, in contrast, are unrelated to changes in behavioral fit.

Of the control variables included in the models, only managerial status and tenure are significant. Managers are more likely to have higher behavioral fit, perhaps because their cultural fit was conducive to their promotion to a managerial position or because subordinates were more likely to accommodate behaviorally to those in positions of authority. Consistent with previous work on enculturation (Srivastava et al., 2017), we also find that individuals exhibit significantly lower behavioral fit during their first year in the organization.\footnote{Tenure has a curvilinear relationship with behavioral fit, steadily increasing during the first six to twelve months and gradually}
Table 4 reports the analyses we used to test our second hypothesis—that being connected to colleagues with higher (lower) behavioral fit will be associated with corresponding increases (decreases) in perceptual accuracy and hence behavioral fit for the focal individual. Model 1 presents estimates from the baseline fixed effect models with lagged peer behavioral fit, as specified in eq. [6] Individuals exhibit a significant increase in behavioral fit when their peers’ mean behavioral fit increases in the preceding month. Importantly, this model includes individual fixed effects and thus accounts for a wide range of time-invariant individual differences—such as self-monitoring or cultural capital—that might also affect a person’s capacity for behavioral fit.

As noted above, the estimates from Model 1 are not causal given that this empirical approach cannot distinguish the effects of homophily, or seeking out similar others, from those of social influence, or modifying one’s own behavior to accommodate others’ behavior. We therefore turn to our instrumental variable in the remaining models. The primary result is reported in Model 2. The coefficient for peer behavioral fit suggests that those who, as a result of the reorganization, transitioned into a network comprising peers with greater behavioral fit experienced an increase in their own behavioral fit in the following month. The opposite is also true: individuals who, through the reorganization, transitioned into a network of peers with lower behavioral fit experienced a corresponding decline in their own behavioral fit. Interestingly, and likely because reorganizations are disruptive to cultural integration, the majority of employees experienced a decline in peer behavioral fit, and correspondingly, their own behavioral fit during this period.

We illustrate the implications of induced change in peer behavioral fit in Figure 4. The diagram plots the effects of the reorganization on individuals’ behavioral fit over time, as estimated by the instrumental variable model. The upper line corresponds to individuals who experienced a half standard deviation positive increase in their peers’ behavioral fit, and the lower line corresponds to individuals who experienced a decline of the same magnitude in their peers’ behavioral fit[11] Both translate to similarly sized adjustments in the stabilizing thereafter. Because individuals vary significantly in speed and steepness of enculturation, we use a binary indicator for early tenure.

[11]These are substantial changes in peer behavioral fit but not implausible during a period of reorganization. A little over 1%
focal individuals’ behavioral fit, but in opposite directions. Moreover, both adjustments persisted for roughly two months, after which the effects of the reorganization were no longer apparent and individuals converged toward mean behavioral fit.\footnote{Because the reorganization was not a true natural experiment, it is worth noting that changes that occurred after its effects were initially felt could have arisen for a variety of reasons that we do not observe in our data. For example, individuals presumably regained more command over whom they interacted with after the reorganization, which would also reintroduce potentially confounding homophily effects. Hence, the period immediately following the reorganization is the appropriate one to consider for this analysis.}

Importantly, the two sets of individuals—positively and negatively “treated”—are indistinguishable in the period preceding the reorganization, suggesting that these adjustments are a result of the imposed change in network composition rather than systematic differences between the two groups. The Kleibergen-Paap F statistic, which is appropriate when using robust standard errors, suggests that the instrument is strong.\cite{Kleibergen2006, Baum2007}

Changes in the number of peers had a more modest impact: those who experienced an increase in the size of their network due to the reorganization experienced declines in behavioral fit. Forced network growth, in other words, is disruptive to cultural integration. The difference between these coefficients in the OLS (Model 1) and instrumental variable (Model 2) models highlights the importance of causal identification in this context. During non-turbulent times, an increase in number of peers is associated with an increase in behavioral fit. Our results suggest, however, that the increase in network size is driven by improved cultural integration, which facilitates seeking out more contacts in the organization, and not the other way around. When changes are forced, however, attending to a growing number of peers whom the focal individual does not necessarily choose to interact with appears to undermine cultural adjustment.

Our models do not speak directly to how precisely this cultural transmission occurs—for example, whether organizational members explicitly reward and penalize their colleagues for culturally compliant or deviant behavior or whether cultural knowledge is transferred tacitly. Models 3 and 4—wherein we estimate the effects of change in peer behavioral fit on the focal individual’s perceptual accuracy and value congruence, respectively—suggest that behavioral adjustment occurs through changes in perceptual accuracy rather than through value congruence. We conjecture that individuals adapt their perceptions, but not their private experiences a positive shock at or greater than half a standard deviation, but roughly 35% experienced a decline of that magnitude.
beliefs, in response to changes in peer composition. Moreover, in Models 5 and 6 we estimate the effects of reorganization-driven changes in peer perceptual accuracy and in peer value congruence on the focal individual’s perceptual accuracy and value congruence, respectively. Both coefficients are insignificant, lending further support to our argument that cultural learning occurs through observing peers’ behaviors, given that cognition is less directly accessible to others. We conjecture that the majority of this cultural transmission happens tacitly. As Models 5 and 6 imply, individuals generally do not have access to their peers’ backstage cultural fit. To the extent that they do, for example, when they explicitly discuss their beliefs, it does not appear to be sufficiently potent to translate into changes in their own backstage cognition.

In Table 5, we report the results of two supplemental analyses designed to assess the robustness of the results of our instrumental variables analysis. First, given that our measures of backstage and frontstage cultural fit are all defined with respect to the reference group of an individual’s interlocutors in a given month, which people can—to varying degrees—self-select into, we replicated the instrumental variables analysis using behavioral fit and peer behavioral fit measures that were based on the reference group of all employees in the organization. Table 5, Model 1, shows that peer behavioral fit, when peers are defined as all other employees in the organization, predicts the focal actor’s behavioral fit relative to this same reference group. This result helps mitigate concerns that our main results are an artifact of our choice to define behavioral fit relative to a focal actor’s interlocutors in a given month.

Second, our instrumental variables approach is predicated on the assumption that the reorganization produced exogenous shifts in focal actors’ peer groups. Yet it is possible that the reorganization was engineered to produce certain desired shifts in peer groups—for example, distancing leaders and their teams when there was animosity between leaders or bringing together formal subunits whose heads had compatible management styles. To address such possibilities, we replicated the analyses using a sub-sample of employees who were not in supervisory roles. We reasoned that, insofar as the reorganization was designed in part to change peer groups, such social engineering was targeted to the leadership ranks of the company. For those in individual contributor—rather than supervisory—roles, the reorganization was much more likely to have produced exogenous change in peer networks. As Table 5, Model 2, illustrates, our hypothesized effects hold even for
this more restricted sample of employees.

[TABLE 5 ABOUT HERE.]

Discussion and Conclusion

Adjustments to new and changing cultural environments are a fixture of modern life. People’s identities in contemporary society typically intersect many social boundaries—including ethnic, religious, political, occupational, and organizational. This crisscrossing of boundaries requires ongoing cognitive and behavioral effort. The contemporary workplace—with its growing emphasis on culture on the one hand and employees’ declining average tenure on the other—is a central stage on which these cultural transitions are enacted. Navigating the cultural heterogeneity across and within organizations involves maintaining multiple and partial commitments to different cultural orders, which in turn requires cultural awareness and adaptability (Morris, Chiu, and Liu 2015; Friedland and Alford 1991; Schröder, Hoey, and Rogers 2016; DiMaggio and Goldberg forthcoming).

Whether in the context of immigration, education, or organizations, existing literature has tended to approach cultural assimilation through the lens of socialization (e.g., Van Maanen and Schein 1979). Such an approach assumes that cultural adaptation entails a gradual internalization of the group’s norms and underlying value system. Research on organizational culture has therefore predominantly examined cultural fit as the match between the individual’s and the group’s values. We acknowledge that value congruence is important for successful integration into an organization; however, our results indicate that it only tells part of the story. Drawing on and adapting Goffman’s distinction between the back- and frontstage (Goffman 1959), we introduced a conceptual framework that understands cultural fit as comprising both value congruence and perceptual accuracy. As our findings demonstrate, both value congruence and perceptual accuracy matter for individual productivity. Yet perceptual accuracy, not value congruence, is associated with an individual’s capacity to behave in normatively compliant ways. That said, we acknowledge that linguistic fit is not the only way for those with high levels of value congruence to display normative compliance. For example, given the robust link between value congruence and longevity found in previous research, it

13That said, we acknowledge that linguistic fit is not the only way for those with high levels of value congruence to display normative compliance. For example, given the robust link between value congruence and longevity found in previous research, it
influence her own perceptual accuracy and hence her ensuing behavioral fit.

Our findings inform sociological research on organizations, culture, and the intersection between the two. Prior work on the link between organizational culture and individual attainment has consistently shown that cultural fit yields positive outcomes for individuals and the organizations they inhabit (Chatman and O’Reilly [2016]). Whereas this work has mostly focused on value congruence as the primary dimension of cultural fit, our study contributes to a burgeoning line of research that explores the more fine-grained behavioral aspects of cultural fit—such as linguistic conformity—and the factors that produce variation in individuals’ behavioral patterns of enculturation (Srivastava et al. [2017]). Shifting the focus from values to perceptions, we also demonstrate that the ability to read the cultural code—what we term “perceptual accuracy”—is inherently related to individuals’ capacity to behave in a culturally compliant manner.

The results from this investigation also add insight into the question of why some people enculturate more successfully than others. Previous work has often assumed that enculturation is a function of individual differences in endowments. Rivera (2012), for example, demonstrates that labor market matching—at least in the elite firms she investigates—is inherently related to the cultural capital that job applicants possess. Separately, research by organizational psychologists has focused on innate differences in psychological traits. It demonstrates that stable dispositions such as self-monitoring and perspective-taking are conducive to cultural adjustment and its positive implications (Maddux, Mullen, and Galinsky [2008]). In contrast, we use an instrumental variable approach to show that the ability to enculturate is also contextual (cf. Ashforth, Sluss, and Saks [2007]), accruing to individuals whose peers are themselves successfully enculturated. Cultural adaptation, in other words, is not just a function of the ability to read the cultural code but also of the peers from whom this code is learned.

Although our findings were drawn from a single organizational setting, raising questions about generalizability, we conjecture that they may apply more broadly. In particular, this work adds to our understanding of cultural dynamics in large social groups. Extending Goffman’s (1959) dramaturgical metaphor, our study raises the partition that typically separates the backstage of cognition from the frontstage seems likely that if a member is not involuntarily separated from the organization, she is likely engaging in certain other behaviors that are normatively compliant.
Lifting the Curtain

of behavior, demonstrating how cognition and behavior are intertwined in producing and sustaining cultural order. Recent work in cultural sociology has distinguished shared preferences from shared meanings and construals (Goldberg, 2011; DiMaggio and Goldberg, forthcoming). Our findings show that the latter—that is, agreement in how a situation is interpreted, not necessarily in what is desirable or worthy—is sufficient for an identifiable culture to emerge. Our distinction between perceptual accuracy and value congruence provides an analytical framework for understanding how cognition and behavior can converge or diverge. Future work might draw on these foundations to further our understanding of how cultural homogeneity and heterogeneity emerge and why, despite cognitive fragmentation at the individual level, culture can nevertheless appear to be coherent to the social group as a whole.

This work also contributes to a growing body of research examining the interrelationships between structure and culture (Goldberg et al., 2016; Lizardo, 2006; McLean, 1998; Srivastava and Banaji, 2011; Hunzaker, 2016). Our results indicate that individuals’ propensity to fit into their groups—both in their thoughts and in their actions—can be influenced by the behavior of group members with whom they interact. Previous work has argued that an innate “cultural intelligence” makes individuals sensitive to cultural knowledge in others’ behaviors (Tomasello, 2009; Henrich, 2016). A sociological literature on social networks, in contrast, has mostly focused on the structural conditions that enable or impede behavioral diffusion. We combine insights from these otherwise disconnected lines of work to make two interrelated contributions. First, we demonstrate that cultural transmission is a function not only of individuals’ attentiveness to cultural knowledge in others’ behaviors (Tomasello, 2009; Henrich, 2016). A sociological literature on social networks, in contrast, has mostly focused on the structural conditions that enable or impede behavioral diffusion. We combine insights from these otherwise disconnected lines of work to make two interrelated contributions. First, we demonstrate that cultural transmission is a function not only of individuals’ attentiveness to cultural knowledge in others’ behaviors, but also to the structural conditions that lead and expose them to others. Second, we show that this process of cultural diffusion operates, first and foremost, by affecting perceptions.

Although sociologists have demonstrated that individuals can be influenced by others to adjust their values and beliefs (Marsden and Friedkin, 1993), our findings add nuance to this understanding by suggesting that peer influence with respect to values may take longer and require more exposure than influence related to perceptions of the normative order. Interesting questions about the mechanisms that underpin the interpersonal transmission of culture naturally follow. For example, how does interacting with someone who is perceptually accurate or behaviorally compliant allow one to correct one’s own misperceptions? When do interactions
Lifting the Curtain

with group members who exhibit high levels of value congruence affect one’s own preferred values? Which network structures might sustain versus disrupt stable behaviors? Answering questions such as these likely requires pinning down underlying causal mechanisms, likely through careful laboratory investigations.

The conceptual separation of backstage fit into value congruence and perceptual accuracy points to several other promising directions for future research. We speculate, for example, that value congruence may provide a motivational channel through which a person is more or less vigilant in achieving and maintaining perceptual accuracy. We similarly conjecture that people with chronically low value congruence may be able to maintain high perceptual accuracy for a finite period of time but that doing so may, over time, adversely affect their identity and sense of self-worth (cf. Hochschild 2012). Conversely, even if those with high perceptual accuracy and low value congruence do not experience intrapsychic conflict, they may still experience the deleterious effects of being judged by others as inauthentic. Alternatively, we speculate that such an individual may—through self-perception and attribution processes (Ross 1977)—begin to experience an increase in value congruence. Questions such as these about the dynamic interplay between the dimensions of the backstage and their long-term consequences for frontstage conformity, individual well-being, and group harmony remain to be explored.

Finally, through this work, we make a methodological contribution that would appear to have wide-ranging application across the social sciences. Building on Salganik’s (2017) notion of “amplified asking,” we demonstrate an empirical approach that transforms a one-time self-report into a longitudinal data set. Such an approach may not be appropriate in all situations. For example, it requires a sufficient number of survey observations, access to rich communication content, protocols and safeguards to protect individual privacy and company confidentiality, and significant computational time to run models. Yet, given the ubiquity of digital trace data, the increasing difficulty of collecting survey data (particularly over time and from a large number of organizations), the widespread dissemination of off-the-shelf machine learning tools, and the declining cost of processing capacity, we anticipate that the pairing of self-reports and digital trace data will become increasingly common in sociological research (Lazer and Radford 2017, Evans and Aceves 2016, McFarland, Lewis, and Goldberg 2016). We see great potential for such work to more fully illuminate
how the frontstage and backstage of social life relate to one another and jointly shape the life course and the cultures in which it unfolds.
References


Herrmann, Patricia A, Cristine H Legare, Paul L Harris, and Harvey Whitehouse. 2013. “Stick to the Script: The Effect of Witnessing Multiple Actors on Children’s Imitation.” *Cognition* 129:536–543.


Ostroff, Cheri Lee and Tim Judge. 2007. Perspectives on Organizational Fit. Psychology Press.


Potts, Christopher. 2011. “Sentiment-Aware Tokenizer.” Creative Commons Attribution-NonCommercial-ShareAlike 3.0 Unported License: http://creativecommons.org/licenses/by-nc-sa/3.0/.


Figures

FIGURE 1: A schematic illustration of our theory. Four individuals (A-D) are each characterized by their values (V), perceptions (P) and behavioral probabilities (B). Arrows correspond to causal relationships.

FIGURE 2: Conceptual Overview of the Machine Learning Process
FIGURE 3: OLS and fixed effect regressions of perceptual accuracy and value congruence, with indicators for each tenure month up to 36 months in the company.

FIGURE 4: Marginal effects, estimated by monthly consecutive instrumental variable models, of change in peer behavioral fit on individual behavioral fit. The two lines correspond to individuals who experienced a 0.5 increase (blue) or decrease (red) in peer behavioral fit. Shaded areas correspond to 95% confidence intervals.
## Tables

**TABLE 1**

**Fixed Effect Regressions of Bonus (logged)**

<table>
<thead>
<tr>
<th></th>
<th>Model 1</th>
<th>Model 2</th>
<th>Model 3</th>
<th>Model 4</th>
</tr>
</thead>
<tbody>
<tr>
<td>Behavioral Fit†</td>
<td>0.131***</td>
<td></td>
<td>0.122***</td>
<td></td>
</tr>
<tr>
<td></td>
<td>(4.45)</td>
<td></td>
<td>(4.14)</td>
<td></td>
</tr>
<tr>
<td>Perceptual Accuracy†</td>
<td>0.144***</td>
<td>0.122**</td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>(3.97)</td>
<td></td>
<td>(3.05)</td>
<td></td>
</tr>
<tr>
<td>Value Congruence†</td>
<td></td>
<td></td>
<td>0.056**</td>
<td>0.046†</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>(3.18)</td>
<td>(2.37)</td>
</tr>
<tr>
<td>Manager</td>
<td>-0.194</td>
<td>0.025</td>
<td>0.063</td>
<td>-0.180</td>
</tr>
<tr>
<td></td>
<td>(-1.12)</td>
<td>(0.13)</td>
<td>(0.31)</td>
<td>(-1.02)</td>
</tr>
<tr>
<td>Constant</td>
<td>5.642***</td>
<td>5.394***</td>
<td>5.299***</td>
<td>5.666***</td>
</tr>
<tr>
<td></td>
<td>(28.18)</td>
<td>(26.63)</td>
<td>(25.68)</td>
<td>(28.47)</td>
</tr>
<tr>
<td>Individual FE</td>
<td>Yes</td>
<td>Yes</td>
<td>Yes</td>
<td>Yes</td>
</tr>
<tr>
<td>Department FE</td>
<td>Yes</td>
<td>Yes</td>
<td>Yes</td>
<td>Yes</td>
</tr>
<tr>
<td>Year FE</td>
<td>Yes</td>
<td>Yes</td>
<td>Yes</td>
<td>Yes</td>
</tr>
<tr>
<td>Observations</td>
<td>4785</td>
<td>6379</td>
<td>6379</td>
<td>4780</td>
</tr>
<tr>
<td>Num. Individuals</td>
<td>1058</td>
<td>1304</td>
<td>1304</td>
<td>1057</td>
</tr>
<tr>
<td>( R^2 )</td>
<td>0.059</td>
<td>0.043</td>
<td>0.040</td>
<td>0.065</td>
</tr>
</tbody>
</table>

† statistics in parentheses; standard errors clustered by individual

† lagged variables, * \( p < 0.05 \), ** \( p < 0.01 \), *** \( p < 0.001 \)
### TABLE 2

**Competing Risks Model of Voluntary Exit**

<table>
<thead>
<tr>
<th></th>
<th>Model 1</th>
<th>Model 2</th>
</tr>
</thead>
<tbody>
<tr>
<td>Perceptual Accuracy</td>
<td>1.005</td>
<td></td>
</tr>
<tr>
<td></td>
<td>(0.07)</td>
<td></td>
</tr>
<tr>
<td>Value Congruence</td>
<td></td>
<td>0.876*</td>
</tr>
<tr>
<td></td>
<td></td>
<td>(-2.30)</td>
</tr>
<tr>
<td>Manager</td>
<td>0.833</td>
<td>0.864</td>
</tr>
<tr>
<td></td>
<td>(-0.77)</td>
<td>(-0.62)</td>
</tr>
<tr>
<td>Female</td>
<td>1.386*</td>
<td>1.392*</td>
</tr>
<tr>
<td></td>
<td>(2.53)</td>
<td>(2.56)</td>
</tr>
<tr>
<td>Age</td>
<td>0.901**</td>
<td>0.902**</td>
</tr>
<tr>
<td></td>
<td>(-3.23)</td>
<td>(-3.23)</td>
</tr>
<tr>
<td>Age$^2$</td>
<td>1.001**</td>
<td>1.001**</td>
</tr>
<tr>
<td></td>
<td>(3.20)</td>
<td>(3.22)</td>
</tr>
<tr>
<td>Num. Employees</td>
<td>1.002***</td>
<td>1.002***</td>
</tr>
<tr>
<td></td>
<td>(9.46)</td>
<td>(9.96)</td>
</tr>
<tr>
<td>Department Dummies</td>
<td>Yes</td>
<td>Yes</td>
</tr>
<tr>
<td>Observations</td>
<td>27467</td>
<td>27467</td>
</tr>
<tr>
<td>$\chi^2$</td>
<td>172.161</td>
<td>177.689</td>
</tr>
<tr>
<td>Log-Likelihood</td>
<td>-1320.27</td>
<td>-1318.36</td>
</tr>
</tbody>
</table>

Exponentiated coefficients; $t$ statistics in parentheses
Standard errors clustered by individual; Sample weights by tenure

* $p < 0.05$, ** $p < 0.01$, *** $p < 0.001$
## TABLE 3

**cross-sectional and longitudinal fixed effects regressions of behavioral fit**

<table>
<thead>
<tr>
<th></th>
<th>Cross-Sectional</th>
<th>Longitudinal</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Model 1†</td>
<td>Model 2†</td>
</tr>
<tr>
<td>Perceptual Accuracy‡</td>
<td>0.122***</td>
<td>0.149***</td>
</tr>
<tr>
<td></td>
<td>(3.56)</td>
<td>(3.37)</td>
</tr>
<tr>
<td>Value Congruence‡</td>
<td>-0.008</td>
<td>-0.040</td>
</tr>
<tr>
<td></td>
<td>(-0.17)</td>
<td>(-0.86)</td>
</tr>
<tr>
<td>Manager</td>
<td>0.613***</td>
<td>0.599***</td>
</tr>
<tr>
<td></td>
<td>(6.73)</td>
<td>(4.20)</td>
</tr>
<tr>
<td>First Year</td>
<td>-0.246**</td>
<td>-0.351***</td>
</tr>
<tr>
<td></td>
<td>(-3.20)</td>
<td>(-3.49)</td>
</tr>
<tr>
<td>Female</td>
<td>0.043</td>
<td>-0.033</td>
</tr>
<tr>
<td></td>
<td>(0.62)</td>
<td>(-0.35)</td>
</tr>
<tr>
<td>Age</td>
<td>-0.003</td>
<td>-0.002</td>
</tr>
<tr>
<td></td>
<td>(-0.84)</td>
<td>(-0.30)</td>
</tr>
<tr>
<td>Constant</td>
<td>0.345*</td>
<td>0.223</td>
</tr>
<tr>
<td></td>
<td>(2.37)</td>
<td>(1.13)</td>
</tr>
<tr>
<td>Individual FE</td>
<td>No</td>
<td>No</td>
</tr>
<tr>
<td>Department FE</td>
<td>Yes</td>
<td>Yes</td>
</tr>
<tr>
<td>Year FE</td>
<td>No</td>
<td>No</td>
</tr>
<tr>
<td>Observations</td>
<td>386</td>
<td>209</td>
</tr>
<tr>
<td>R²</td>
<td>0.275</td>
<td>0.235</td>
</tr>
</tbody>
</table>

* t statistics in parentheses; standard errors clustered by individual when individual fixed effects are used
† Behavioral Fit is averaged over 3 months, ‡ Imputed measures in Models 4-6

* p < 0.05, ** p < 0.01, *** p < 0.001
## TABLE 4

### OLS and Instrumental Variables Fixed Effects Regressions of Behavioral Fit

<table>
<thead>
<tr>
<th></th>
<th>OLS</th>
<th>Instrumental Variable</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Model 1</td>
<td>Model 2</td>
</tr>
<tr>
<td>Peer Behavioral</td>
<td>0.221***</td>
<td>0.266***</td>
</tr>
<tr>
<td>Fit†</td>
<td>(12.68)</td>
<td>(6.38)</td>
</tr>
<tr>
<td>Peer Perceptual</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Accuracy†</td>
<td>(0.63)</td>
<td></td>
</tr>
<tr>
<td>Peer Value</td>
<td>0.001**</td>
<td>-0.013*</td>
</tr>
<tr>
<td>Congruence†</td>
<td>(3.11)</td>
<td>(-2.50)</td>
</tr>
<tr>
<td>Num. Peers†</td>
<td>0.365***</td>
<td>0.555***</td>
</tr>
<tr>
<td></td>
<td>(7.67)</td>
<td>(4.34)</td>
</tr>
<tr>
<td>Manager</td>
<td>-0.154***</td>
<td>-0.204***</td>
</tr>
<tr>
<td></td>
<td>(-6.72)</td>
<td>(-4.12)</td>
</tr>
<tr>
<td>First Year</td>
<td>-0.065</td>
<td>0.648**</td>
</tr>
<tr>
<td></td>
<td>(-1.23)</td>
<td>(2.67)</td>
</tr>
<tr>
<td>Constant</td>
<td>Yes</td>
<td>Yes</td>
</tr>
<tr>
<td>Individual FE</td>
<td>Yes</td>
<td>Yes</td>
</tr>
<tr>
<td>Department FE</td>
<td>Yes</td>
<td>Yes</td>
</tr>
<tr>
<td>Year FE</td>
<td>Yes</td>
<td>Yes</td>
</tr>
<tr>
<td>N</td>
<td>22080</td>
<td>21998</td>
</tr>
<tr>
<td>Num. Individuals</td>
<td>1515</td>
<td>1508</td>
</tr>
<tr>
<td>R²</td>
<td>0.28</td>
<td></td>
</tr>
<tr>
<td>Kleibergen-Paap F</td>
<td>8.99</td>
<td>8.99</td>
</tr>
</tbody>
</table>

† $t$ statistics in parentheses; standard errors clustered by individual

† lagged variables, instrumented endogenous variables in Models 2-6

** $p < 0.01$, *** $p < 0.001$
TABLE 5: Robustness Checks—Instrumental Variables Fixed Effect Regressions of Behavioral Fit

<table>
<thead>
<tr>
<th></th>
<th>Model 1</th>
<th>Model 2</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Peer Behavioral Fit</strong></td>
<td>0.235***</td>
<td></td>
</tr>
<tr>
<td>Organization</td>
<td></td>
<td></td>
</tr>
<tr>
<td><strong>Peer Behavioral Fit</strong></td>
<td>0.158***</td>
<td></td>
</tr>
<tr>
<td>(Organization)</td>
<td>(5.40)</td>
<td></td>
</tr>
<tr>
<td>Num. Peers †</td>
<td>-0.003</td>
<td>-0.013*</td>
</tr>
<tr>
<td></td>
<td>(-1.85)</td>
<td>(-2.10)</td>
</tr>
<tr>
<td>Manager</td>
<td>0.133***</td>
<td></td>
</tr>
<tr>
<td></td>
<td>(3.57)</td>
<td></td>
</tr>
<tr>
<td>First Year</td>
<td>-0.034*</td>
<td>-0.150**</td>
</tr>
<tr>
<td></td>
<td>(-2.27)</td>
<td>(-3.25)</td>
</tr>
<tr>
<td>Constant</td>
<td>2.154***</td>
<td>-0.560</td>
</tr>
<tr>
<td></td>
<td>(26.90)</td>
<td>(-1.79)</td>
</tr>
<tr>
<td>Individual FE</td>
<td>Yes</td>
<td>Yes</td>
</tr>
<tr>
<td>Department FE</td>
<td>Yes</td>
<td>Yes</td>
</tr>
<tr>
<td>Year FE</td>
<td>Yes</td>
<td>Yes</td>
</tr>
<tr>
<td><strong>N</strong></td>
<td>19938</td>
<td>18097</td>
</tr>
<tr>
<td>Num. Individuals</td>
<td>1229</td>
<td>1257</td>
</tr>
<tr>
<td>Kleibergen-Paap F</td>
<td>3.03</td>
<td>8.81</td>
</tr>
</tbody>
</table>

† statistics in parentheses; standard errors clustered by individual
†† instrumented and lagged endogenous variables
* p < 0.05, ** p < 0.01, *** p < 0.001
Appendix A: Machine Learning Procedure

Overview

The procedure consisted of five major steps, which are documented at a conceptual level in Figure 2 in the main manuscript and described in greater detail below.\[\textsuperscript{14}\]

Our first step was to translate the raw email data into a format that is usable by the random forest model. We tokenized and stemmed all words in the body of email messages. Tokenization involves separating the text into distinct terms, for which we used the TwitterTokenizer designed for linguistic analysis Potts (2011). Stemming involves reducing each term to a root form, for which we used the Porter Stemmer from the python nltk package. We removed all characters that could not be encoded into unicode, such as “\x00,” and split the text into n-stems, where \(n\) is in the set \([1,2,3]\). Given that language use tends to follow the power law, in which few terms are used frequently and many terms are used infrequently, we then undertook steps to reduce the dimensionality of the data to make it computationally tractable. We retained all n-stems in emails sent from individuals, but only uni-stems in emails sent to individuals. Additionally, we retained only those n-stems that were used by 99% of employees in a subsample of emails. Finally, we used principal component analysis (PCA) to further reduce dimensionality, retaining only the top 3,000 PCA components for each type of n-stem. These resulting components served as the feature inputs to our model.

The second step was to transform our measures of backstage cultural fit into categories that are more conducive to classification given the relatively small number of observations from which we had to fit the model. Recall that perceptual accuracy and value congruence were computed as correlations, ranging from 0 to 1. We transformed these continuous measures into three discrete categories—low, medium, and high. Intuitively, this allowed our model to detect distinctive features of belonging to each category, an important characteristic to which we will return when we discuss the testing of our model. For perceptual accuracy, we set the cutoffs for low fit at 20% and for high fit at 80%, with everything else considered medium fit. For value congruence, for which we had even fewer observations, we had to set more extreme cutoffs at 10% and

\[\textsuperscript{14}\text{Requests for source code can be sent to the corresponding author. We use python’s \texttt{sklearn} module to implement our machine learning model.}\]
90% to achieve strong model fit.

The third step was to use our feature inputs and their now-discrete mappings to backstage cultural fit to train a random forest model. The random forest model is an ensemble method, which means it aggregates and blends multiple independent decision trees (Ho 1995; Friedman et al. 2001). After several such decisions according to specific features of the input, all of the inputs are sorted into decision leaves. The random forest model then collects those independent trees and their leaves and predicts results for new observations. New observations get sorted into resultant leaves depending on their own features, and their probabilities of being predicted as a certain class depend on the other data points sorted into that leaf in the trained model. In a simplistic model, imagine that the only decision is that $\text{PCA1} > 0.5$ and that all observations with $\text{PCA1} > 0.5$ are high in cultural fit. Then, a new observation whose $\text{PCA1} > 0.5$ would also get sorted into the same leaf and would then be classified as high cultural fit.

The fourth step was to evaluate the trained model. To do so, we assessed the model’s predictions compared to the original continuous values. Random forest models produce, along with the classifications of input, probabilities of observations belonging to each class. Conceptually, this means that if an observation has certain characteristics that correspond to a given class, it will have a higher probability of being in that class. For example, if an individual’s email communication has indicators of low, medium, and high backstage cultural fit, but more indicators of high cultural fit than the others, then his or her output from the random forest model might indicate a 0.2 probability of low fit, a 0.3 probability of medium fit, and a 0.5 probability of high fit. We can then take a weighted sum of these probabilities to generate a measure that is conceptually analogous to the original continuous measure. We used a mix of methods to evaluate the model, including the area under the curve of the receiving operating characteristic curve (ROC AUC), precision-recall, and separation between low and high backstage cultural fit with respect to the original continuous values. As reported in Appendix B, the final models we used performed well on these evaluations.

The final step was to impute perceptual accuracy and value congruence using their corresponding 

---

15A decision tree is an algorithm that divides the input according to distinct decision points. For example, if the algorithm detects that half the input has a first principal component value of greater than .5 and the other half a value of less than or equal to .5, then the decision would be $\text{PCA1} > .5$. 

46
random forest models for all individuals in all time periods for which we had corresponding email data. To do this, we followed the first step above to retrieve the input feature vector for each individual over time and used all the linguistic data for each individual up to a certain month to impute perceptual accuracy and value congruence for that individual in that month.

There were a total of over five million unique emails. Each email can be sent from an individual and several other individuals (via the to/cc/bcc lines). We included both messages sent to and received from the focal individual in our final model.

**Dimensionality Reduction of Features**

Considering the size of our potential feature vector, we used dimensionality reduction techniques to make our process computational tractable. In particular, we used a discriminative heuristic to determine which n-stems to keep, since there is a tradeoff between keeping frequent and non-frequent terms: frequent terms allow for discrimination to the extent that they are used differently among a large population of people, while non-frequent terms allow for discrimination to the extent that some people use them and others do not. Given this trade-off, we retained those n-stems that were used by at least 99% of all employees, regardless of their objective frequency. To retain as much information from this pared down set of n-stems, we used principal component analysis (PCA). This allowed us to reduce the hundreds of thousands of features to only a few thousand per n-stem, while still retaining a large part of the variance of the original data. Because of the exponential size of the “to” stems compared to the “from” stems, we ended up using the top 3,000 PCA components from the “from” uni-, bi-, and tri-stems, and from the “to” uni-stems.

**Random Forest Model Specification**

We selected the random forest model because of several favorable characteristics. First, random forest models allow for nonlinear relationships between input and output. Decision trees in general, of which random forest is a collection, thus allow for arbitrarily complex relationships, which we would assume govern the relationship between linguistic data and cognitive cultural fit. Second, random forests are ensembles of
decision trees, which inherently reduce overfitting and increase robustness. Since there is the potential for a link between linguistic data and cognitive cultural fit to be extremely idiosyncratic (e.g., use of a certain phrase or way of communicating), it greatly helps that we use a more robust method. Third, random forest models do not require as much training data as neural networks. Deep neural networks have the same, if not better, ability to pick up complex relationships, but require far more training data, depending on the depth of the model. As a result, random forest models are simpler and tend to require fewer training data for comparable results.

We split the data into the usual training, development, and testing sets, with 56% of the original data in the training set, 14% in the development set, and 30% in the testing set. Because of the way the random forest algorithm is implemented, it is strongly vulnerable to the “class imbalance” problem. Specifically, if the input to the model from the training set were 10% class 0, 80% class 1, and 10% class 2, then the model would err towards predicting most new observations as class 1. To overcome this, we used a bootstrapping procedure that randomly samples with replacement the lesser classes until they reach the amount of the most populated class. This procedure ensured that, on average, input classes were balanced and therefore class prediction depended more on the splits than on the original balance of the input classes. In addition to searching the hyperparameter space, we also tested varying N for bootstrapped samples.

Appendix B: Evaluating Model Fit

Test Set Metrics

Because of the way we constructed our pseudo-continuous imputed cultural fit, we needed to use a set of test metrics that accurately capture what it means to have a “good model.” The choice of bounds for the continuous to discrete distribution is forced; it is an educated guess that produces empirically validated results. Therefore, observations that lie just on one side may not differ substantively from observations that lie just on another side. Concretely, observations that are on the high end of the medium cultural fit may be very similar to observations that are on the low end of the high cultural fit, given that we had set the cutoff
Lifting the Curtain

ourselves. Therefore, our measures should focus less on perfect categorization (i.e., precision, recall), and more on separation of low and high cultural fit and predictive power of imputed results on actual results. As a result, our performance metrics are a mix of the traditional machine learning metrics, as well as novel metrics we developed ourselves.

For the traditional test metrics, we present the pairwise precision and recall measures on the test set. We provide the pairwise precision recall rather than an F score, because we differentially care about the pairwise results. That is, we care the most about the precision recall between the high and the low cultural fits and less about the precision recall between the mid and either high or low cultural fits, as per our previous discussion.

[TABLE B1 ABOUT HERE.]

A better metric might be to directly examine the separation between groups. If we link the original continuous values with the classifications, then we would see a split like this.

[FIGURE B1 ABOUT HERE.]

We then used the means and standard deviations of each group to see if the classifier successfully split the observations into statistically distinct groups. We find that the separation between low and high in our models is good.

[TABLE B2 ABOUT HERE.]

Finally, we used the receiver operating characteristic curve (ROC) that has become popular in machine learning. Since the ROC works with threshold probabilities of classification, mapping the true positive rate versus the false positive rate at different thresholds, it conceptually measures the extent to which the rank-ordering of predicted values is in line with expectations. For a perfect area under the curve (AUC), the rank-ordering would be monotonically increasing such that all actual values of 1 would have higher
probabilities of being classified as 1 than all actual values of 0, and vice versa. Since we have three classes versus the regular binary classification, we use the micro-averaged ROC curve, which takes into account this structure. The ROC curves with their AUC’s are presented below.

[TABLE B3 ABOUT HERE.]
Appendix B Figures

FIGURE B1: Division of Continuous Cultural Fit into Classes
Appendix B Tables
### TABLE B1

**Test Set Precision-Recall Metrics for Imputations**

<table>
<thead>
<tr>
<th></th>
<th>Precision Low-High</th>
<th>Precision Low-Mid</th>
<th>Precision Mid-High</th>
<th>Recall Low-High</th>
<th>Recall Low-Mid</th>
<th>Recall Mid-High</th>
</tr>
</thead>
<tbody>
<tr>
<td>PA-Interloc.</td>
<td>0.857</td>
<td>0.726</td>
<td>0.767</td>
<td>0.267</td>
<td>0.651</td>
<td>0.711</td>
</tr>
<tr>
<td>PA-Org.</td>
<td>1</td>
<td>0.875</td>
<td>0.865</td>
<td>0.547</td>
<td>0.867</td>
<td>0.849</td>
</tr>
<tr>
<td>VC-Interloc.</td>
<td>1</td>
<td>0.952</td>
<td>0.950</td>
<td>0.667</td>
<td>0.952</td>
<td>0.934</td>
</tr>
<tr>
<td>VC-Org.</td>
<td>1</td>
<td>0.923</td>
<td>0.951</td>
<td>0.667</td>
<td>0.923</td>
<td>0.906</td>
</tr>
</tbody>
</table>
TABLE B2

P-VALUES FOR DIFFERENCE IN MEANS BETWEEN LOW AND HIGH

<table>
<thead>
<tr>
<th></th>
<th>P-Value</th>
</tr>
</thead>
<tbody>
<tr>
<td>PA-Interloc.</td>
<td>2.661e−3</td>
</tr>
<tr>
<td>PA-Org.</td>
<td>1.874e−8</td>
</tr>
<tr>
<td>VC-Interloc.</td>
<td>8.500e−6</td>
</tr>
<tr>
<td>VC-Org.</td>
<td>7.157e−5</td>
</tr>
</tbody>
</table>

TABLE B3

AREAS UNDER THE ROC CURVE

<table>
<thead>
<tr>
<th></th>
<th>ROC AUC</th>
</tr>
</thead>
<tbody>
<tr>
<td>PA-Interloc.</td>
<td>0.740</td>
</tr>
<tr>
<td>PA-Org.</td>
<td>0.910</td>
</tr>
<tr>
<td>VC-Interloc.</td>
<td>0.950</td>
</tr>
<tr>
<td>VC-Org.</td>
<td>0.930</td>
</tr>
</tbody>
</table>