Sociological Forum

Sociological Forum, Vol. 36, No. S1, December 2021
DOI: 10.1111/socf.12769

Revision, Reclassification, and Refrigerators[1]

Terence E. McDonnell,[2] Dustin S. Stoltz,[3] and Marshall A. Taylor[4]

Current debates about cultural change question how and how often change in personal culture happens. Is personal culture stable, or under constant revision through interaction with the environment? While recent empirical work finds attitudes are remarkably stable, this paper argues that typifications—how material tokens are classified as a particular mental type by individuals—are more open to transformation as a result of the fundamentally fuzzy nature of classifying. Specifically, this paper investigates the social conditions that lead people to reclassify. How do we move people to see the same thing differently over time? Paying attention to type–token dynamics provides mechanisms for why and under what circumstances personal culture may change. To assess reclassification, the paper analyzes an online survey experiment that asked people to classify refrigerators as owned by "Trump" or "Biden" voters. Those participants who received definitive feedback about the correct answer were more likely to reclassify than are those receiving normative feedback about how "most people" classified the images. Implications for cultural change and persuasion are discussed.

KEYWORDS: classification; cognition; culture; mechanisms; online survey experiment; type–token dynamics.

INTRODUCTION

Classification is essential to interpretation and action. Classifying objects in the environment is entwined with acting within that environment, because, as Gestalt theorists and ecological psychologists argued, "[objects'] looks tell us what to do with them" (Chemero 2003; Gibson 2014; Koffka 1935:76). How we "clump" things together develops implicitly over repeated experiences interacting with objects in the environment, as well as interacting with or distinguishing ourselves from others on the basis of those classifications (Chemero 2003; Lizardo and Strand 2010; Mandler 2004; Martin 2011). Through such interactions, people develop relatively entrenched categories which operate via principles of prototypicality and schematicity as a result of prioritizing efficiency over accuracy (Koriat et al. 2000; Martin 2010). People do not use exhaustive rule-based operations of inclusion and exclusion. Instead, people predominantly sort objects *gradationally* by fuzzy family resemblances (Cerulo 2006;

[1] We would like to thank the University of Virginia's Sociological Workshop on Aesthetics, Meaning, and Power (especially Fiona Rose Greenland, Isaac Reed, Adam Slez, and Natalie Aviles), the University of Notre Dame's Culture Workshop (especially David Gibson and Ann Mische) for their excellent feedback and suggestions. This paper has also benefited from conversations with Erin Metz McDonnell, Jennifer Lena, and Michael Wood. Extra special thanks to special issue editors Karen Cerulo and Vanina Leschziner for their leadership.
[2] Department of Sociology, University of Notre Dame, 4060 Jenkins Nanovic Hall, Notre Dame, IN, USA; e-mail: terence.e.mcdonnell@nd.edu
[3] Department of Sociology and Anthropology, Lehigh University, Bethlehem, PA, USA; e-mail: dss219@lehigh.edu
[4] Department of Sociology, New Mexico State University, 292B Science Hall, Las Cruces, NM, USA; e-mail: mtaylor2@nmsu.edu

Lakoff 2008; Rosch 1999). Classification, then, can be messy. The world is not so crisp as to be neatly classified without revision.

Despite the potential messiness of the world, much work on culture and cognition finds personal culture becomes "locked in" and relatively stable over time (Bourdieu 1977; Gartman 2007; Kiley and Vaisey 2020; Vaisey and Kiley 2021; Vaisey and Lizardo 2016). Such stability is reinforced by bias and selective perception in the face of incongruent information (Bourdieu 1988; DiMaggio 1997; Hunzaker 2016; Strand and Lizardo 2017). In spite of the inherent difficulties of changing people's settled sense of the world, however, organizations pour time and resources into actively persuading others to revise their personal culture (McDonnell 2016). Public health organizations alone annually spend millions of dollars on anti-smoking (cigarettes from cool to cancer), HIV-prevention (condoms from awkward to sexy), and vaccine promotion (risky to safe) campaigns. Reading recent culture and cognition research would lead one to conclude that such persuasion efforts are a Sisyphean task. If Martin (2015:249) is right that "[m]ost of what we think of as cultural change is really change in the composition of populations," are attempts to significantly change personal culture worth the effort?

For example, consider efforts to retrain police to remove implicit racial bias so they do not easily classify Black men as criminals (FitzGerald et al. 2019; Smith 2015). The literature on implicit bias and policing suggests that for many U.S. police officers the category of "dangerous" is associated with skin color (Kahn and Davies 2011; Kahn et al. 2016). Can the personal culture of individual police officers be revised, or must cohorts of police officers be replaced to facilitate change in the organizational culture more broadly? If personal culture is truly stable and durable, maybe we should give up any hope of persuading changes in personal culture around implicit racial bias. Alternatively, rapid changes in public opinion about same-sex marriage and the legalization of marijuana suggest it may be too soon to give up on such efforts.

This paper asks *how do we move people to see the same thing differently over time?* We take the position that *type–token dynamics* provide feasible mechanisms for personal culture change. *Types* are cultural objects in the head, therefore "mental, learned, and instantiated as cognitive processes" (Taylor et al. 2019:2). *Tokens* are the tangible material manifestations of cultural objects out in the world, perceptible through the senses, and with qualities that exist independently of an individual's perception (ibid). For example, you may have a pencil "type" in your head, a general association of expected features and material qualities, while if you look in the drawer in your desk you may find varied pencil "tokens": long red pencils, short yellow pencils with a fat pink eraser, mechanical pencils, colored pencils, etc. You cannot have types without tokens, and vice versa. Given their independence, though, tokens vary in ways that imperfectly map onto types. Taylor et al. (2019), therefore, suggest that the binding of material tokens to mental types is open to revision. Through exposure to new material tokens, people reflect upon cognitive types, sometimes wrestling with unintended and unexpected qualities of new tokens. These dynamics of mismatch, we argue, are sites of destabilization and change.

Type–token dynamics may lead to distinct outcomes. First, as people are exposed to more token objects, it is intuitive to presume that types will be reinforced.

Reinforcement is when a person takes an established type and applies it to an object, often ignoring qualities of the object that are not prototypical (Bourdieu 1984:175; Lizardo 2004:386–7). However, *variations in the qualities* of those objects may lead to type revision and some variation is inevitable in even the most faithful replications (Taylor et al. 2019:11). By expanding the type to include these new qualities (Lizardo 2004:387–8), the type may require maintenance—that is, change in the service of stability (Taylor et al. 2019:14). Alternatively, a token or set of tokens may be *reclassified* as an altogether different type. It is reclassification that is the subject of this paper. Under what circumstances is a type–token binding affirmed and stabilized, versus questioned, reclassified, and revised? Are there social conditions that make people more likely to reclassify?

To assess reclassification, we used an online survey experiment that repeatedly exposed people to repeated similar tokens, asked them to classify the tokens into one of two types, and then later asked them to classify some of the same tokens a second time. We randomly sorted participants into different feedback conditions to assess what social contexts are more likely to push people to reclassify. One of the treatment groups was *normatively* influenced by giving participants information about how "most people" had classified the objects, while another treatment group was *definitively* told the right or wrong answer. Each treatment was compared to a control group that received no feedback. We picked a task in which classification is ambiguous: the token could be reasonably classified as either of two types. This was important because we wanted a baseline assessment of how much movement was possible for a task where the stakes were low and the type–token relationship was comprehensible but not overdetermined.

As a case, we adapted a *New York Times' Upshot* quiz that asked people to classify refrigerators as owned by "Trump" or "Biden" voters.[5] Participants were shown photographs of the interior contents of refrigerators and then asked to classify the image as the refrigerator of a "Trump" or "Biden" supporter. In effect, we are asking people to extend general types well-developed during an election year (Trump v. Biden voter) to incorporate previously unrelated tokens (refrigerators). Before and after the treatment phase, participants classify six identical images. When participants switch their classification of a refrigerator after the treatment, we define this as reclassification. We then assessed whether our treatment conditions motivated participants to reclassify refrigerators at a higher rate than the control group. Ultimately, we find that those in the *definitive* condition were more likely to reclassify.

This paper moves the field of culture and cognition in new directions by pushing cultural sociologists to isolate the conditions under which personal culture changes. In what follows we articulate how work on types, tokens, reclassification, and revision can answer new questions about the links between culture and cognition. After presenting our findings, we then discuss the implications of these results for understanding reclassification, persuasion, and personal culture change. We conclude by pointing toward future research that could come from this agenda.

[5] https://www.nytimes.com/interactive/2020/10/27/upshot/biden-trump-poll-quiz.html

THEORETICAL FRAMEWORK

Classifying and Cultural Change

Culture and cognition research is now a dominant approach within cultural sociology, and with good reason. Dual process approaches upended assumptions about how culture works by attending to how people *process* information and acquire personal culture (Cerulo 2010; DiMaggio 1997; Leschziner and Brett 2019; Leschziner and Green 2013; Lizardo et al. 2016; Vaisey 2009). Scholarship on the senses has revisited and updated our understanding of how perception and attention shape interpretation of the environment (Cerulo 2018; Friedman 2011; Lembo 2020; Rose-Greenland 2016). The intersection of culture and cognition has also drawn attention to the ways "public culture" and "personal culture" are in dynamic tension (Cerulo 2018; Engman and Cranford 2016; Lizardo 2017; Luft 2020; Nolan 2020; Strauss and Quinn 1997; Wood et al. 2018). Debates about cultural change underlie these agendas. How can moments of deliberative processing reconfigure personal culture? How do meanings change or stabilize through people's sensory interactions with the material environment? How does public culture "out there" reshape personal culture "in here" and vice versa?

Kiley and Vaisey (2020) recently articulated the contours of this debate about cultural change by drawing out two models. The "active updating" model suggests that as people interact with their environment, they continuously reshape their dispositions to better fit environmental information. The "settled dispositions" model suggests that early experiences stabilize personal culture when people are young. Once settled, interactions with the environment and new information do little to change those dispositions. The active updating model, by contrast, sees personal cultural change as ongoing as people confront unfamiliar and novel objects in the environment. The settled dispositions model sees broader change at the cohort level—a new generation with different but stable dispositions replace older generations, leading to public cultural transformation. Their study of three waves from the General Social Survey offers evidence more consistent with the settled dispositions model, and for those items "that do show evidence of active updating, the overall rate of persisting change in the population is likely low . . . [for most items] less than 1 percent of the population appears to make any persisting change" (Kiley and Vaisey 2020:20).

Kiley and Vaisey's study casts doubt on whether active updating happens much at all or leads to durable change, although they leave open space for some active updating related to public or peripheral topics.[6] This is discouraging news for sociologists who see in culture the power to persuade individuals to change how they see and act in the world. One adjacent literature at the intersection of culture and

[6] While we broadly agree with Kiley and Vaisey's assessment of this debate, we worry they may not have a complete picture. Their 2020 paper lacks a sufficient measure of "environment." Is the stability of people's personal culture happening because personal culture overrides the environment, or is it because people's environments are so routinized and stable that their dispositions are rarely challenged? We suspect less routinized backgrounds may encourage more active updating than can be assessed with the General Social Survey.

cognition may offer a compromise between settled dispositions and active updating: *classification*.

Studying classification through a cognitive lens is renascent in cultural sociology. “Lumping and splitting” (Zerubavel 1996) is both a cognitive and a cultural process. Building on classic work in anthropology (Douglas [1966] 2003; Durkheim [1912] 1995:141–150; Durkheim and Mauss [1963] 2009; Lévi-Strauss 1963; Sahlins 1976; Turner 1970) and recent work in cognitive psychology and linguistics (Lakoff 2008; Margolis and Laurence 1999; Rosch 1973), there is resounding agreement that classic “rule-based” categorizing—for example, bachelors are unmarried men, therefore all unmarried men are bachelors—is relatively rare (Margolis 1994; Taylor 2003), if not impossible in the final instance (Bourdieu 1977:22–30; Garfinkel 1967:26–7). Instead, tokens are bound to a type via prototypicality, “scalar goodness-of-example judgments for categories” (Lakoff 2008:136).

Rather than either-or operations of inclusion and exclusion, people predominantly sort objects gradationally by family resemblances (Rosch and Mervis 1975; Zadeh 1996). Importantly, “to speak of a prototype at all is simply a convenient grammatical fiction; what is really referred to are judgements of degree of prototypicality” (Rosch 1999:200). In addition to prototypicality, the acquisition of types operates via schematization. Schematizing may reinforce already existing types by extracting the patterns most redundant and most in agreement with one’s prior experiences, supporting a view of people primarily as “meaning maintainers” (Patterson 2004, 2013). However, schematization also entails systematic forgetting of “specific details of a new experience and retain only the generalities that liken the event to other experiences in one’s past” (Cerulo 2002:8; see also Wood et al. 2018). Taken together, prototypicality and schematization reveal mechanisms by which classifications, as fuzzy sets, are *open to revision* (Stoltz and Taylor 2017; Taylor et al. 2019).

Table I. Varieties of Type–Token Processes.

		Is the type–token relation changed?
	Ignoring Tokens match types, incongruent qualities of tokens are ignored	No
Revision	***Maintaining*** Tokens do not match types, types are expanded to accommodate incongruent qualities	**No**
Revision	***Reclassifying*** Tokens previously classified under an established type later become classified under a different established type	**Yes**
	Innovating Existing types cannot account for new token, either because of the novelty of the token or social feedback, and a new type is formed	Yes

Bolding highlights the kinds of binding “revision” that may take place, and on which the present article focuses.

Largely obscured in debates about the resistance or propensity for personal cultural change is the question of *what* changes. While values and attitudes may be stable, type–token dynamics and classifications may be more subject to change given the fuzzy relationship between them. Specifically, we call attention to processes of type *revision* and *reclassification*. Revision refers to how individuals alter their existing types and reclassification is when a token first sorted into one type is later sorted into another. Much of the early classification literature omits the distinction between the personal and public by assuming a nearly one-to-one correspondence, and thereby obfuscating any possible misalignment between types and tokens. To the extent research on classification has considered type dynamics at the level of personal culture, it has mostly focused on the acquisition and emergence of new types. Shepherd (2010), for example, has shown how the internalization of classification systems at the level of personal culture is shaped by, but also partially decoupled from, external classification systems. Guilbeault et al. (2021) explore how people converge on shared ways of classifying ambiguous tokens (cf Hebart et al. 2020), specifically how groups of people come to share the same types over multiple iterations of coordination, in line with research on the development and maintenance of "arbitrary traditions" (e.g., Jacobs and Campbell 1961; Zucker 1977). In this paper, by extension, we consider how established types may change.

Revision and Reclassification

How do individuals come to revise their classifications? As people engage the world they make sense of their environment through classification (or typification), accruing the contours of cognitive types from repeated experience with the regularities of material tokens. Classifications are stabilized when embodied anticipations continually "predict" the affordances of material tokens: our expectations are continually met. Expectations, however, are not entirely an individual process—they necessarily involve coordination with others (Sutton et al. 2010). Similarities in classifications provide a scaffolding for social action, and therefore we continually attend to the feedback provided by both the physical world pushing back on us and others' externalizing their own judgments in their actions and words.

We argue that attention to type–token dynamics reveals a set of possible pathways to change. Let us consider some possible type–token processes (see Table I). If tokens in the environment do not change, there is little reason for types to adapt—regardless of how many tokens we encounter, they will always reinforce preexisting types. As much work in cognition suggests, in the face of new tokens people's instantiated types are plenty good enough to make decisions in the world, even if those decisions are biased or incomplete. Confirmation bias, for example, leads people to treat new tokens as confirmation of existing types, *ignoring* the ways that those new tokens might be different or might challenge those types. Therefore, it is reasonable to expect that new tokens, even if not precisely the same, usually stabilize to existing types. Pushing beyond stability, we can theorize three kinds of type–token dynamics which entail change in personal culture.

Taylor et al. (2019) theorized two processes emerging from type–token dynamics that might produce change: innovating and indexicalizing. *Innovating* happens when a new token resists classification, either resulting from the token's sheer novelty or feedback from others, and a new type is formed to make sense of the token. *Indexicalizing* is the process by which a token is identified as an instance of type, produced as an instantiation of a type, or alters a type to make room for that token. Through the unavoidable variation in the qualities of tokens—even in the most faithful reproductions— people may need to *revise* a preexisting type through expansion or contraction of the prototypical qualities of the type.

In this paper, we offer a third process emerging from type–token dynamics: *reclassifying*. Reclassifying occurs when people come to view a token, previously classified under an established type, as now under a different, yet established type. For example, how things once classified as "bad" become read as "good" or items associated with the "outgroup" become associated with the "ingroup." Importantly, we argue that reclassifying may lead to type revision. *Revision* is a likely outcome of repeated reclassification because regularly moving tokens from one category to another casts doubt on the validity of those classifications. *Reclassifying* is the process we seek to understand in this paper, specifying the social contexts in which reclassifying is mostly likely to occur.

Social Contexts of Typification

For some cultural objects, especially the most generic or basic, the binding of a token to a type is largely motivated by the redundant and invariant features of our world (Gibson 2014; Mandler 2008; Martin 2011). As we move to more complex types, however, social context is likely to play a greater role (Martin 2011). At the most basic, we may use language, not to describe or name, but to simply guide one's attention to certain qualities of a token (Mandler 2008:224–6). Finally, we may also use language to challenge or affirm other's typification by providing *feedback* about the correctness or propriety of their interpretations.

To motivate our hypotheses, we draw on research about performance, norm persistence, and conforming to others' perceptions in small groups. Broadly literature on performance finds that feedback leads to a larger impact than, for example, punishment and rewards, and furthermore, the way feedback is provided is significant (Hattie and Timperley 2007). Similarly, research on persistence finds that the context of communication is an important predictor of consistency over time (Zucker 1977). Much of the literature on feedback is directly linked to some measure of accuracy (Morris and Moore 2000). We, however, are interested in a step often in the service of accuracy: change. Therefore, we also draw on literature explicitly probing the question of *persistence*, or the lack of change in the so-called "arbitrary traditions" (Jacobs and Campbell 1961; Zucker 1977). Finally, as we are interested in how individuals may change previous typifications, we build on work on high-order inferences, which explores the interplay between one's own judgment and "most people's" judgment (Correll et al. 2017; Firat et al. 2018; Melamed et al. 2020; Ridgeway and Correll 2006). Drawing on this work, we hypothesize that whether feedback leads to reclassifying depends upon the level of consensus conveyed: is the feedback definitive or merely normative.

When feedback is *definitive*, the type to which a token belongs is presented as a certainty. In Zucker's (1977) study of cultural persistence, participants were asked to estimate how much a small light directed at a dark wall moved. Due to the autokinetic effect, the participants perceive the light moving, when in fact it is fixed. When alone and receiving no feedback, there is a large variance in estimates among participants, but this standard remains consistent for the individual (Sherif 1935). In transmission studies using the autokinetic effect (Jacobs and Campbell 1961), a "naive" participant is informed of a previously established "group standard" by a confederate of the researchers. The confederate represents someone already "enculturated," as if they are passing down a tradition. Regardless of the ratio of enculturated to naive participants, the estimates converge with the control group's standard after the fourth or fifth generation—in other words, the "tradition" decays. Zucker (1977) found that greater "institutionalization," or "taken-for-grantedness" of the situation, leads participants to change more toward the confederate's estimate, greater stability over several generations, and more stability and resistance to change a week later. Institutionalization, according to Zucker, "reflected certainty in the accuracy of the group standard." *We therefore hypothesize that when feedback is presented as definitive, participants will be more likely to reclassify a token than a control group.*

Conversely, when feedback is normative, the type to which a token belongs is presented as a matter of public opinion, albeit the majority's opinion. A consistent finding is that "what most people think" can shape and even override what a person personally thinks (Firat et al. 2018; Ridgeway and Correll 2006; Sobotka 2020). In a study of status advantage, for example, Correll et al. (2017) asked people to rate and recommend a series of chocolates in a blind taste test. Although the chocolates were actually the same, the participants are informed at various points that one chocolate is a high-status brand and another is a low-status brand. Believing status conveys what "most people" would think is the superior chocolate, the participants' inferences about others' perceptions override their personal perceptions. Indeed, even when told that the recipient of the chocolate preferred the low-status brand, participants were still more likely to recommend the high-status brand. This was also the case when status information did not resolve uncertainty. Aligning one's own beliefs with others' perceptions is also supported by Garfinkel's classic "breaching" experiments in which incongruences with member's expectations prompted attempts to "restore a right state of affairs" (Garfinkel 1967:42). If normative claims are contrary to the participant's classification, this is not simply a matter of personal preferences or motivations, but rather suggests they have missed something that most people have not (Garfinkel 1967:56–7). Therefore, *we hypothesize that when feedback is presented as normative, participants will be more likely to reclassify than a control group.*

RECLASSIFICATION IN AN AMBIGUOUS TASK

For the present study on reclassifying, we compare differences between a baseline test and a re-test. Participants are presented with a token and asked to classify it as being more prototypical to one of two types. The participants are randomly assigned to the control group or treatment conditions designed to influence reclassifying. Our primary outcome, then, is the total number of tokens reclassified.

Unlike previous work on change in personal culture, we are not looking at the stability over time, but rather across contexts. Furthermore, unlike research on the classification of ambiguous objects, we are not looking at how a person may classify a previously unknown object. Rather, we focus on creating an ambiguous task, in which it is the *binding* between a type and a token that is very likely to be novel for the individual. It is important to note that this is a first step into reclassifying and future work should consider the durability of reclassifying over time and the social context of maintenance. Here we attempt to isolate the social context of feedback that will create uncertainty or certainty about a person's initial classification.

The Case

To explore the social context of reclassification, we use the case of the *New York Times'* Upshot story on how people classify refrigerators as "Trump" or "Biden" voter's refrigerators.[7] Using photographs of the interior contents of people's refrigerators, participants are asked to sort them as a "Trump" or "Biden" refrigerator. More precisely, participants attempt to assimilate a new index (i.e., refrigerator) to a preexisting type (i.e., "right-leaning person" or "left-leaning person").

This is a case of mapping political ideology to consumption practices. It may be possible that participants are drawing on stereotypical or empirical associations between brands or food items and a person's given ideological inclination. Indeed, there is empirical support for the notion that some lifestyle choices are highly correlated with political attitudes despite having no necessary connection to politics (DellaPosta et al. 2015). The present task, however, was selected precisely because accurately classifying most of the images appears to be nearly random. In the *New York Times* poll of over 3 million guesses, participants were accurate just 52% of the time. In the aggregate, accurately classifying a refrigerator as belonging to either a "Trump" or "Biden" supporter is as likely as flipping a coin.

We intentionally sought a case where classifying is ambiguous. We leverage ambiguity because it would likely be difficult to move people to reclassify tight and well-established type–token pairs in a short survey.[8] The New York Times refrigerator quiz is an ideal test because the types used to classify tokens are clear, timely, and salient, but the tokens they must extend to were open to interpretation. This openness to interpretation is facilitated by the numerous objects in each refrigerator, and the multiple ways those objects could be interpreted. These refrigerators are rich with information; participants could look at the kinds of foods, brands, cleanliness of the refrigerator, orderliness of the refrigerator, the quality of containers, the amount of organic or sugary items, the cost of items, the rarity of goods, or more.

[7] https://www.nytimes.com/interactive/2020/10/27/upshot/biden-trump-poll-quiz.html
[8] Although certainly this is a motivating assumption that is open to empirical scrutiny.

To offer a sense of what participants would confront when seeing an image, here is a description of one refrigerator in our sample. This image was typical for the sample.

> Top Shelf (left to right): one half-gallon carton of milk, two half-gallon containers of orange juice (one Florida Natural cardboard cartons, one plastic bottle that looks like Simply Orange) sit next to a reusable plastic container of what looks like homemade chicken soup. From there is a block of cream cheese (Lucerne), a container of Feta, and what appears to be a bag of shredded white cheese.
>
> Middle Shelf (left to right): glass container with what could be cooked chicken breasts, covered in cellophane; a square package of raw ground beef (grass fed if you read closely), a bag of sliced deli turkey with a plastic container with a red lid of cooked rice sitting on top of the turkey, and next to that a plastic container with a blue top containing something creamy inside.
>
> Bottom Shelf (left to right): a half-used package of raw bacon sits atop a clear plastic carton of a dozen eggs, which sits atop an 18 egg container made from styrofoam. Three jars of pickles sit to the right, at least one has sliced hamburger pickles. To the right is a glass bottle of Coca-Cola, two LaCroix cans of different flavors, with two white beverage cans of different sizes (impossible to tell their contents). To the right of the beverages is what looks like a bag of shredded yellow cheese, with some herbs (parsley or cilantro?) leaning against it.
>
> Much of what is in the back of the refrigerator appears to be leftovers from past meals. The refrigerator is mostly full but not overstuffed; organized by content, but not orderly. This refrigerator was a Biden supporter's refrigerator.

DATA AND METHODS

Sample and Participants

The sample consisted of participants from a Qualtrics panel. The data collection was overseen and managed by a project manager assigned by Qualtrics. Participants were surveyed in February 2020. Participants who completed the survey in under half the median time to completion were not retained, leading to a final sample size of 777 participants. All participants are U.S. citizens and at least 18 years of age. Demographic characteristics for the sample are provided in Table II. The sample moderately skews female (57%), Democrat (43%), and suburban (36%). The sample is heavily skewed white (79%). These skews are to be expected with online samples without demographic quotas (Buhrmester et al. 2011; Casler et al. 2013). Since all participants were randomly assigned to a treatment or control condition—detailed in the "Design" section below—the distribution of participants per condition is fairly uniformly distributed across the levels of these covariates. Participants were compensated for their participation. The median time to completion for the entire survey was 11.12 minutes. The survey consisted of a survey experiment and a sociodemo-

Table II. Sample Descriptives (N = 777).

	Total	Mean (SD)
Gender/Sex		
Male	326	
Female	443	
Genderqueer/Non-Conforming	5	
Transgender Male/Transgender Man	2	
Other	1	
Race/Ethnicity		
White	612	
Hispanic	139	
Black or African American	443	
Asian	40	
American Indian or Alaska native	4	
Native Hawaiian or other Pacific Islander	1	
Black or African American, White	4	
American Indian or Alaska Native, Asian, White	1	
American Indian or Alaska Native, White, Other	1	
Other	18	
Political orientation		
Republican	255	
Democrat	332	
Independent	177	
Other	13	
Location		
Big city	184	
Mid- or Small-size city	128	
Suburbs	275	
Small town/Rural	189	
Other	4	
Experimental condition		
Control	125	
Affirmed	129	
Definitive/Actual	129	
Definitive/Inverted	131	
Normative/Actual	129	
Normative/Inverted	134	
Reclassification		1.865 (1.468)
Baseline accuracy		2.997 (1.023)
Recognize more than 50% of items		0.259 (0.438)
Interested in politics		0.723 (0.448)

graphic questionnaire. For the experimental portion, the research design (see Fig. 1) contained five training conditions and a control.

Design

Image Sets: We divided images of refrigerators into four sets. Set A contained six images that participants encountered in both the baseline test and the re-test. It is Set A that we use for our outcome variable of reclassification (which we discuss in more detail in the "Variables" section below). Importantly, participants did not see Set A during the training conditions discussed below. Set B contained an additional six images that participants encountered in the baseline test. Participants randomly allo-

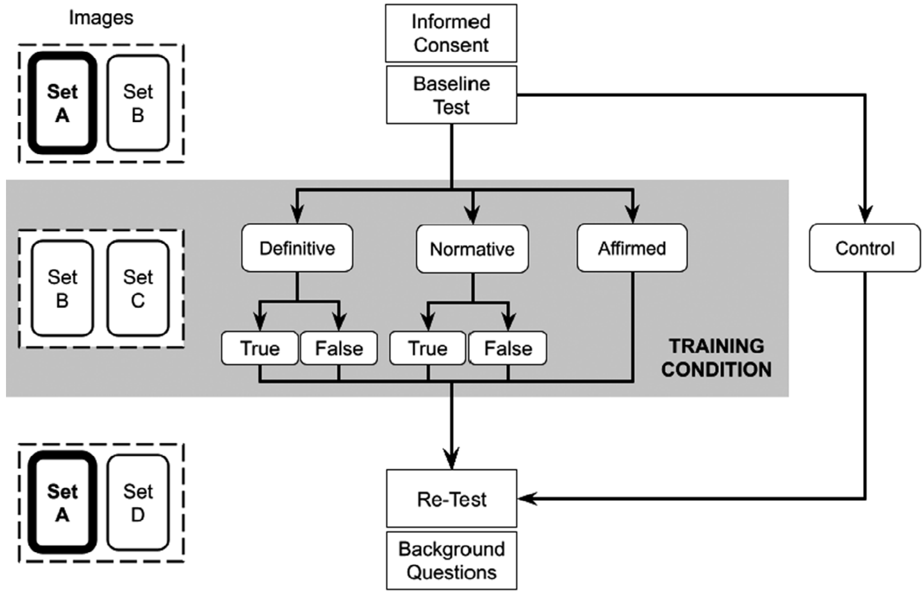

Fig. 1. Diagram of Study Design.

cated to a training condition saw Set B again, were reminded of their initial classification, and received feedback about the accuracy of their classification as discussed below. Participants in the training conditions also saw an additional six images, Set C, which they had not already classified. Finally, during the re-test, participants saw Set A again along with an additional six images, Set D, that they had not already seen. Each set contained three images of Biden supporter refrigerators and three images of Trump supporter refrigerators.

Baseline: After providing informed consent, participants completed a baseline test by indicating whether they believed a given refrigerator belonged to either a Trump supporter or Biden supporter. In all, 12 such images were classified in this manner and each participant classified the same 12 images (i.e., Set A and Set B). Participants were asked to answer as quickly as possible.

Treatment and Control: Our primary hypotheses tested the relative effects of receiving *definitive* and *normative* feedback. We also tested the effects of facticity of the feedback and receiving affirming feedback resulting in six possible conditions: (1) Definitive/Actual; (2) Definitive/Inverted; (3) Normative/Actual; (4) Normative/Inverted; (5) Affirmed; and (6) Control (see Table III). We discuss each below.

> *Definitive Condition:* Individuals may be more or less likely to reclassify depending on the *certainty* of feedback. During the training phase, participants in the Definitive condition were told in a matter-of-fact way the type of supporter to whom the refrigerator actually belonged. For example, a participant who indicated a refrigerator belonged to a Trump supporter in the baseline test but

which actually belonged to a Biden supporter saw the following statement for that image in the Definitive condition: "You said this was a Trump refrigerator. This refrigerator belonged to a Biden supporter." We hypothesize that definitive contexts will increase the likelihood of reclassifying.

Normative Condition: We alter the *certainty* of feedback by providing participants with an indicator of "public opinion" about a classification. During the training phase, participants in the Normative conditions were told "how most survey takers" classified that image. However, this did not actually reflect survey takers' answers. Rather, those in the Normative condition were told that most survey takers chose the actual candidate supported by the owner of the refrigerator. For instance, a participant who indicated a refrigerator belonged to a Trump supporter but which actually belonged to a Biden supporter saw the following statement if they were in the Normative condition: "You said this was a Trump refrigerator. Most survey takers said this refrigerator belonged to a Biden supporter." We hypothesize that normative contexts will also increase the likelihood of reclassifying, but to a lesser extent than the definitive context.

Facticity: In both the Definitive and Normative conditions, we also tested whether giving participants Actual or Inverted feedback would lead to more reclassifying. In the Actual feedback condition, participants would be told the actual answer. When a refrigerator was owned by a Trump supporter, they would be told it was owned by a Trump supporter; if Biden, they'd be told Biden. In the Inverted feedback condition, the response was simply inverted: Trump supporters' refrigerators were represented as belonging to *Biden* supporters, and vice versa. In both conditions, participants were also told their own answer from the baseline. Testing facticity allows us to assess whether getting better, factual feedback has effects independent from the certainty of feedback (Definitive or Normative). In general, actual feedback about their accuracy on a given task may lead people to seriously reconsider their own classification. Indeed, much educational practice and theories of learning are based on this notion (Metcalfe 2017). The actual feedback may forefront qualities of the token which are prototypical of the correct type, and therefore be perceived as more believable or trustworthy information. However, given the ambiguity of our case and near-random association between token and type, we hypothesize that facility will have no effect on reclassifying.

Affirmed Condition: Finally, we test whether reclassifying is instigated by feedback in general, rather than the specific content (i.e., the assessment of one's inaccuracy at the task). We *affirm* participants' answers by stating that a given image is, as a matter of fact, classified as whatever answer the participant gave. Learning that one is "correct" is simultaneously a reference to the "objective" facts and also a positive assessment of the individual, which may reinforce one's original answers (Eil and Rao 2011; Peeters 1971), and therefore stabilize an individual's type. In general, people focus more on "bad news" over "good news" (Baumeister et al. 2001) *unless the feedback is about themselves*, in which case the asymmetry is reversed. In fact, when drawing on declarative memories

Table III. A description of the six experimental conditions.

Condition	Description	Feedback
Definitive/Actual	Participant is told the actual answer for the image (Biden) in a definitive way	"You said this was a Trump refrigerator. This refrigerator belonged to a Biden supporter."
Definitive/Inverted	Participant is told the inverted answer for the image (Trump) in a definitive way	"You said this was a Trump refrigerator. This refrigerator belonged to a Trump supporter."
Normative/Actual	Participant is told the actual answer for the image (Biden) in a normative way	"You said this was a Trump refrigerator. Most survey takers said this refrigerator belonged to a Biden supporter."
Normative/Inverted	Participant is told the inverted answer for the image (Trump) in a normative way	"You said this was a Trump refrigerator. Most survey takers said this refrigerator belonged to a Trump supporter."
Affirmed	Regardless of the actual image, the participant is told their baseline answer (Trump) in a definitive way, affirming they were right. Every image in the training phase from the baseline is affirmed	"You said this was a Trump refrigerator. This refrigerator belonged to a Trump supporter."
Control	No training condition	No feedback

For the purposes of demonstration, we take a situation in which the actual image is a Biden supporter's refrigerator, and in the baseline the participant evaluates it as a Trump Supporters Refrigerator.

from one's own lives, people have a tendency to "erase" bad episodes (Baumeister et al. 2001:343–4; Sedikides and Skowronski 2020; Taylor 1991). We therefore hypothesize that affirming participants' answers will make them more certain about their prior responses and lead to the least reclassification.

Control Condition: To provide a baseline of comparison, participants in the Control condition went straight from the baseline test to the re-test (described below).

In addition to seeing six of the images they had seen in the baseline (Set B), participants in all conditions but the Control were shown six new images (Set C) and were either correctly told the type of supporter to whom the refrigerator belonged (for those in the Definitive/Actual and Affirmed conditions); incorrectly told the type of supporter to whom the refrigerator belonged (Definitive/Inverted); told how "most survey takers" answered, with the answer being the real supporter (Normative/Actual); or told how "most survey takers" answered, with the answer being the wrong type of supporter (Normative/Inverted). The six new images were the same across all participants. Altogether, 12 images were seen during the training conditions—six old (Set C) and six new (Set D). To ensure that participants were engaged during the training, they were asked to click on the spot in the refrigerator image that best indicated that the refrigerator belonged to a Biden/Trump supporter (with the answer depending on whether they were in the Actual or Inverted condition).

Re-test: Finally, participants were re-tested with another round of images, half of which were also included in the baseline test (Set A).[9] Importantly, these six images from the baseline were the six *not reused* in the training condition and therefore they were not informed of the "correct" supporter for the images in Set A. Participants then completed another classification task identical to the baseline. Six new images were also introduced in the re-test (Set D), making for 12 total image classifications in the re-test. Standard demographic characteristics were then collected with a questionnaire after the re-test, with items for gender, race, socioeconomic status, political orientation and interest, education, and urbanicity. We also asked questions about the items that participants recognized in the refrigerator and the types of things they focused on when making their classified decisions.

Variables

Dependent Variable—Reclassification: Our dependent variable is a count of the number of times a participant switched their answer from the baseline to the re-test. To do this, we took the images in Set A—that is, the six images that were seen in the baseline and re-test (but not in the training phase)—and counted the number of times the answer changed: either from "Biden" to "Trump" or "Trump" to "Biden." The lowest possible value was 0 (indicating no reclassification) and the highest possible value was 6 (indicating 100% reclassification).

Independent Variable—Training Condition: Our independent variable is a categorical measure indicating the six training conditions to which the participant was randomly assigned: (1) Definitive/Actual; (2) Definitive/Inverted; (3) Normative/Actual; (4) Normative/Inverted; (5) Affirmed; and (6) Control (i.e., no training). Some of the models reported in the results below use a version of this variable with the facticity dimensions collapsed, that is, Definitive, Normative, Affirmed, and Control.

Control Variables: We control for factors that might affect reclassification and which are endogenous to the experimental task but independent of the treatment conditions.

Although accuracy is low in the aggregate, participants in the Definitive/Actual, Normative/Actual, and Affirmed conditions might be less likely to change their answers in the re-test if their baseline accuracy was high. This is because these participants are more likely to be told that their baseline predictions for Set B were correct,

[9] An "articulation" task also followed both the baseline and re-test tasks, wherein participants were asked to click on the three portions of each image that best indicated whether the refrigerator belonged to a Trump supporter or Biden supporter. These data are not used in this paper and therefore not discussed further.

which might further entrench their preexisting types and make them less likely to reclassify images in Set A. We therefore control for baseline accuracy with a count of the number of baseline images correctly classified, which ranged from 0 (no correct answers) to 6 (all correct answers).

Participants might also be less likely to change their answers in the re-test if they feel particularly "familiar" with the objects in the images. A sense of familiarity may be primed by the baseline task: as they classify, participants might believe they have enough referential anchors in the images to select what they believe to be the right answers and may therefore be unresponsive to the training conditions. In other words, their minds might be "made up" before they get to the training.

We account for familiarity in two ways. First, we control for item recognition. Participants were asked "of the brands in the refrigerators you saw, what percentage of these brands did you recognize?" Responses were converted to a dichotomous variable, where 0 = up to 50% and 1 = more than 50%. Second, to account for political familiarity, participants were asked how interested they were in politics, with answers from "Not at all interested" to "Very interested." We converted responses to a dichotomous variable, where 0 = not interested and 1 = interested.

Due to random assignment, sociodemographic variables were near-uniformly distributed across the conditions (see Table II above). Nevertheless, we did run models with other possible controls, none of which fundamentally altered the results of the models presented.

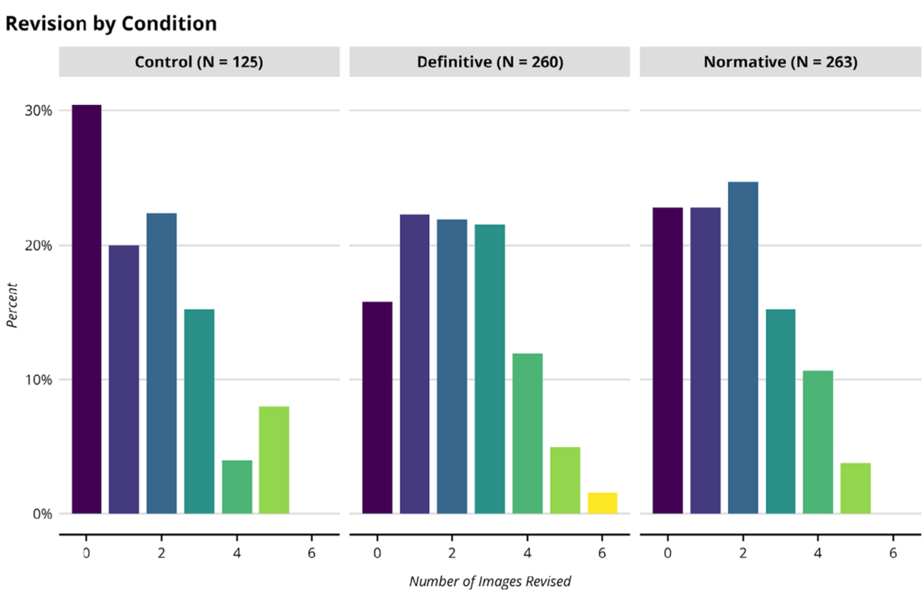

Fig. 2. Percentage of people in each condition by the number of images reclassified Bar colors are mapped to the number of reclassifications. Darker colors/shades mean fewer reclassifications.

Table IV. Hypotheses.

Condition	Hypothesis	Support
Primary Hypotheses		
Definitive	Definitive feedback will lead to more reclassifications than the control condition	+
Normative	Normative feedback will lead to more reclassifications than the control condition	–
Secondary Hypotheses		
Facticity	The facticity of feedback will have no effect on reclassifications	Mixed
Affirmed	Affirmative feedback will not result in more reclassifications than the control condition	+

Analytical Strategy

Reclassification was modeled using OLS regression. Poisson regression models provide nearly identical results and are available upon request.

FINDINGS

In this section, we describe our findings from our online survey experiment. First, we present the results of the test of our Definitive and Normative conditions against the Control condition. Then, we discuss results from the tests of our secondary hypotheses of Facticity and Affirmed. Then, we offer a substantive summary of our results. For a summary of our hypotheses and the extent to which each was supported, see Table IV.

Descriptive statistics of the raw reclassification counts in Fig. 2 reveal some clear patterns. Looking to how many people in each group reclassified *none* of the images, it is evident that the Control group has about twice as many participants who reclassified none of the images (about 30%) than the Definitive condition (about 15%), while the Normative condition lands somewhere in the middle (about 23%). Furthermore, Definitive was the only condition in which there were people who reclassified all six images, and about 40% in the Definitive group reclassified three or more images.[10] Are these differences significant?

Definitive and Normative Conditions

We hypothesized that both the Definitive (i.e., participants read "This refrigerator belonged to a Biden supporter") and Normative (i.e., participants read "Most survey takers said this refrigerator belonged to a Biden supporter") conditions would lead to more reclassifications than the control (no feedback) group. Ultimately, we

[10] Only 5 of 777 participants (0.6%) reclassified all six of the images. As a robustness check, we ran OLS and Poisson models with a six-category version of the dependent variable (ranging from 0 reclassifications to 5/6 reclassifications). There was almost no difference with the models reported in the body of the paper.

Table V. OLS Estimates of Reclassification.

	Model 1		Model 2		Model 3	
	$\widehat{\beta}$	SE_{β}	$\widehat{\beta}$	SE_{β}	$\widehat{\beta}$	SE_{β}
Model 1 treatments (ref = control)						
Definitive	0.451	0.160**				
Normative	0.087	0.161				
Model 2 treatments (ref = control)						
Definitive/Actual			0.575	0.184**		
Definitive/Inverted			0.328	0.184		
Normative/Actual			0.046	0.186		
Normative/Inverted			0.126	0.183		
Model 3 treatments (ref = control)						
Definitive					0.450	0.160**
Affirmed					−0.016	0.185
Controls						
Baseline accuracy	0.128	0.057*	0.130	0.057*	0.106	0.064
Item recognition (ref = "up to 50%")	−0.316	0.133*	−0.311	0.133*	−0.316	0.150*
Political interest (ref = "not interested")	−0.061	0.129	−0.056	0.129	0.127	0.147
Constant	1.429	0.226***	1.419	0.226***	1.352	0.243***
R^2	0.032		0.035		0.037	
RMSE	1.459		1.457		1.461	
$F\,(df_1, df_2)$	4.237 (5, 642)***		3.318 (7, 640)**		3.945 (5, 508)**	
N	648		648		514	

Note: Each model controls for participants' baseline accuracy (i.e., how many images they correctly classified as Trump or Biden supporter refrigerators in the first round), item recognition (i.e., whether participants report recognizing 50% of the items in the image), and participants' degree of political interest.
*p < .05, **p < .01, ***p < .001 (two-tailed tests).

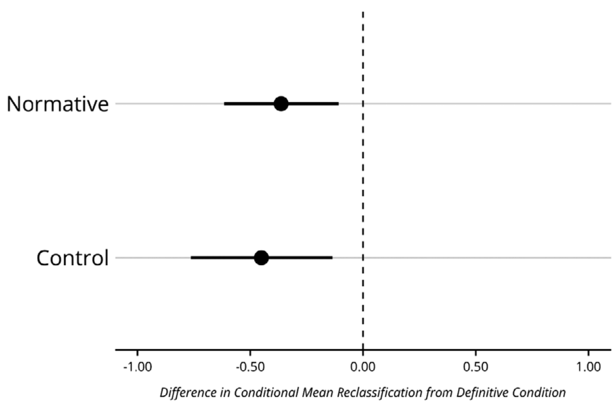

Fig. 3. Marginal Effects for Model 1.
Note: OLS estimates reported. The results are derived from Model 1 in Table IV, which is additive, meaning that these marginal effects are equivalent to each dummy variable's contrast with the definitive condition as the reference category. Any gray bars indicate a statistically non-significant contrast at $\alpha = 0.05$ (two-tailed tests) since the 95% confidence intervals include zero. Control variables are included in the models but not visualized. $N = 648$.

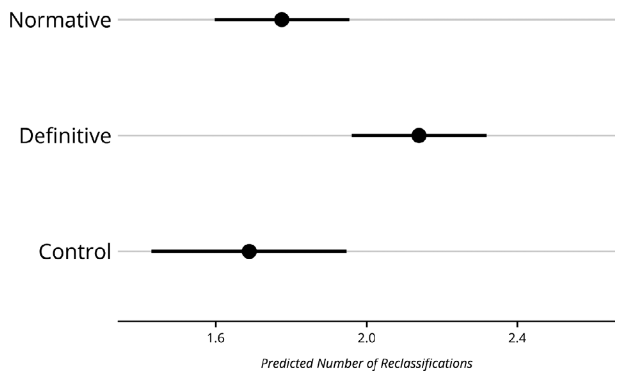

Fig. 4. Adjusted Condition Predictions for Model 1.
Note: Predictions are adjusted predictions at means (APMs), meaning the control variables are all held at their means (Williams 2012). Error bars are 95% confidence intervals.

found support for our hypothesis in the Definitive condition but did not find support for the Normative condition. We present these results in the first model in Table V: it reports the differences in conditional means (conditioning on baseline accuracy, item recognition, and political interest) for the Definitive and Normative treatment conditions relative to the Control.

According to Model 1 (Table V), the Definitive condition is associated with more reclassification as the difference in conditional means between the Definitive and Control conditions is both positive and statistically significant.[11] Participants in the Definitive treatment average about 0.451 more reclassifications than participants in the Control condition. The Normative condition is not significantly different from the Control condition. Figure 3 reports the marginal effects of the Definitive condition. These plots show that participants in the Definitive condition average more statistically significant reclassifications than their counterparts in the Normative condition.[12] Taken together, those people who receive definitive feedback on their baseline classification task average more reclassifications in the re-test task than people who receive no feedback and people who receive normative feedback.

Figure 4 reports adjusted predictions at means (Williams 2012) for the main model (Model 1), which compares the Definitive and Normative conditions against the Control. A participant in the Definitive condition made about 2.14 reclassifications (±0.18, 95% confidence intervals). This is for a participant with mean values across the control variables and regardless of facticity. As Fig. 4 demonstrates, participants in this condition are much more likely to change which political candidate the owner of the refrigerator supported than the Normative or Control conditions.

[11] $\beta_{Def\text{-}Control} = 0.451$; $t = 2.82$; $p = .005$.
[12] $\beta_{Norm\text{-}Def} = 0.364$; $t = 2.81$; $p = .005$.

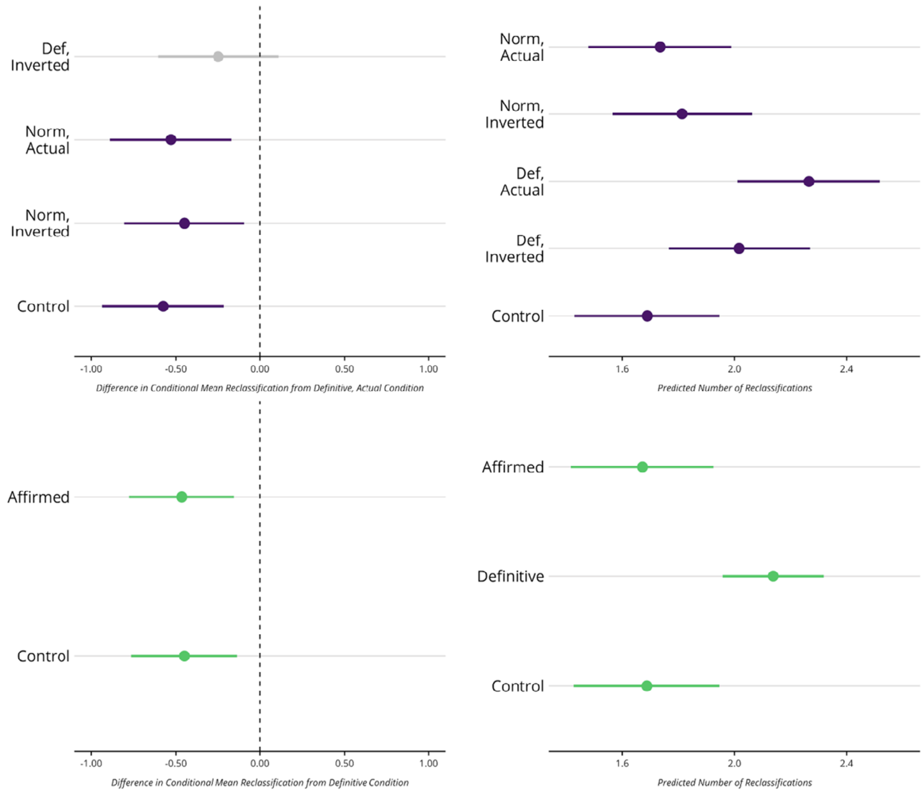

Fig. 5. Marginal Effects and Adjusted Predictions for Models 2 and 3.
Note: Marginal effects and adjusted predictions were calculated the same way as those for Model 1. N = 648 (Model 2, top panels) and N = 514 (Model 3, bottom panels).

Facticity

We sought to assess whether the *facticity of the feedback* had effects upon the likelihood of reclassification, independent of definitive and normative feedback. Given that the classification accuracy by participants was like a coin-flip, we hypothesized that whether a participant receives Actual or Inverted feedback will be unlikely to affect reclassification in this case.

To our surprise, actual feedback did encourage more reclassifying than inverted feedback in the definitive condition (note, however, that this condition had no impact on their accuracy). This finding can be seen in Model 2, which shows differences in conditional means, relative to the Control, for the Definitive and Normative treatments separated out by the facticity sub-conditions (see Table IV). According to the model, the Definitive/Actual treatment has a statistically significant contrast with the Control: all else equal, participants in the Definitive/Actual condition average about 0.575 more reclassifications than participants in the Control condition.[13]

[13] $\beta_{Def/Actual\text{-}Control} = 0.575$; $t = 3.12$; $p = .002$.

The Definitive/Inverted condition tells a similar story, but this difference in conditional means is only significant at $\alpha = 0.1$.[14] This offers support for the idea that feedback matters more when it is true to reality, but that the definitive presentation of the feedback has strong effects *independent of facticity*.

Participants in the Definitive/Actual treatment average more reclassification than both the Normative/Actual and Normative/Inverted conditions (see the top-left panel of Fig. 5 for a plot of the marginal effects of the Definitive/Actual condition).[15] None of the marginal effects of the Definitive/Inverted condition are statistically significant at standard α-thresholds (and therefore not plotted), with the exception of $\widehat{\beta}_{\boldsymbol{Def/Inverted-Control}}$ ($p = .075$). However, like the Definitive/Actual effects, all of the contrasts show the Definitive/Inverted condition with a higher average amount of reclassifying. These results refine our Model 1 results. Namely, while definitive feedback is associated with reclassification over and above what we would expect with normative feedback or no feedback at all, receiving *factual* definitive feedback encourages more reclassification than receiving *false* definitive feedback.

The top panel of Fig. 5 shows the predictions for the second model, which tested the extent that accurate feedback would interact with receiving the feedback in a definitive or normative manner. When separating out the Definitive conditions by facticity, this same participant in the Definitive/Actual condition made about 2.26 reclassifications (±0.25) while a participant in the Definitive/Inverted condition made about 2.02 reclassifications (±0.25). In other words, participants were somewhat more likely to reclassify when getting definitive feedback about the actual political candidate the owner of the refrigerator supported.

Affirmation

Using a definitive affirmed treatment (e.g., participants definitively told they were always correct), we can test another dimension of how language and feedback facticity shape reclassification. We hypothesized that being affirmed you were correct would be no different from the control condition of no feedback, and we found support for this hypothesis. As can be seen in Model 3, being told one is right discourages participants from reflection and reconsideration of their typing. Model 3 is similar to Model 1 but substitutes in the Affirmed v. Control contrast in place of the Normative v. Control contrast (see Table IV). As hypothesized, the Affirmed v. Control contrast is not statistically significant.[16] The Affirmed v. Definitive contrast is statistically significant. As the marginal effects in the bottom-left panel in Fig. 5 shows, participants in the Definitive condition average more reclassification than participants in the Affirmed condition.[17] The bottom-right panel of Fig. 5 shows the predictions for the third model, which test the extent the relationship between receiving definitive feedback and reclassification is mediated by being told one is wrong at least some of the time. As hypothesized, receiving definitive feedback that entirely affirms one's prior classifications tended to reclassify at nearly the same level as the

[14] $\beta_{Def/Inverted\text{-}Control} = 0.328$; $t = 1.79$; $p = .075$.
[15] $\beta_{Norm/Actual\text{-}Def/Actual} = 0.529$, $t = 2.88$, $p = .004$; $\beta_{Norm/Inverted\text{-}Def/Actual} = 0.450$, $t = 2.48$, $p = .013$.
[16] $\beta_{Affirmed\text{-}Control} = 0.016$; $t = 0.08$; $p = .933$.
[17] $\beta_{Affirmed\text{-}Def} = 0.465$; $t = 2.93$; $p = .003$.

Control condition. This finding further complements the main findings from Model 1: definitive feedback leads to more reclassifying than any other treatment/control condition. However, if a participant receives definitive feedback that entirely reaffirms one's prior classifications, they are unlikely to reclassify and are thus entrenching their typifications.

DISCUSSION

What does it take to move people to reclassify? We found that definitive feedback moved participants to reclassify images from the baseline test more often (on average) than the control condition, whereas the normative condition did not have a significant effect compared to the control. In addition, definitive feedback motivated reclassification more than the normative and affirmed treatments. Being told you are always right (affirmed treatment) suppresses reclassification. Put simply, people do not like being wrong, therefore motivating them to reconsider and reclassify tokens; but they do not mind being unpopular if they think they are right.[18]

In the control group about 30% of the participants made no changes, but in the definitive group it was only about 15%. The Definitive/Actual condition group reclassified an average of 2.3 of the six images, compared to 1.7 for the control group. Given how rarely people reclassified 5–6 images (less than 10% in each condition), the magnitude of this nudge is significant. This suggests that to move people toward rethinking their classifications you would want specific conditions: (1) clear and definitive language communicating (2) accurate feedback making clear that (3) some of their past type–token associations were wrong. As people are exposed to new tokens, we infer that these conditions lead people to reflect on those types and consider reclassification.

We were surprised that the Normative condition was not significantly different from the Control group. Claims to the majority—as in "Most survey takers said..."—did not seem to move people to reclassify. Given that literatures on peer-pressure, bandwagon effects, and conforming to majority and group norms we expected people would be moved to reclassify. It is *especially surprising* given the low-stakes conditions of classifying refrigerators.

We suspect that there may be something about the groupist politics of this particular moment that may lead people to stick to their guns in the absence of definitive information. We surmise that it may be easy to dismiss other people's assessments as "opinion," instead of seriously considering the normative assessment and reflecting on their classifications. There has been much talk about former President Trump's use of "Many people are saying"[19] and "A lot of people tell me"[20] and the effectiveness of this rhetoric in selling the veracity of his claims. In light of our findings, such strategies may work to affirm what people already believe but may not be effective at *moving* people off their existing typifications. For those already inclined to align with

[18] Thanks to David Gibson for this pithy summary.
[19] https://www.vox.com/2016/8/18/12423688/donald-trump-speech-style-explained-by-linguists
[20] https://www.theatlantic.com/magazine/archive/2018/03/how-to-talk-trump/550934/

the former President, normative appeals may help bolster people's confidence in pre-existing classification tendencies and encourage bias.

What do these findings tell us about reclassification? For one, it confirms that social conditions have different capacities to influence people to switch their position. Research on confirmation bias suggests that in the face of evidence that contradicts a person's way of seeing the world, people are more likely to perceive that evidence in ways that confirm their worldview. If personal culture remains largely a steady state—as suggested by much work on culture and cognition, most recently by Kiley and Vaisey (2020)—this is likely through mechanisms such as selective perception, confirmation bias, homophily, and the routine interactions made possible by a relatively steady environment. While tokens confirm types in many, if not most situations, this paper has sought to understand what it takes to move people off their typifications. As such, these findings make important contributions to work on cultural change and persuasion.

Implications for the Study of Cultural Change

We argue a focus on reclassification opens up new avenues into debates about cultural change. Reclassification is a form of personal cultural change that does not require changes to attitudes. Say, for example, a person believes that illicit drugs are harmful, subvert the social order, and should be criminalized. People could maintain this settled disposition while reclassifying marijuana, for example, as not an illicit drug but pain relief for cancer patients or belonging to the same type as caffeine and alcohol. These acts of reclassification allow people to maintain their attitudes, generally, while fomenting change at the level of personal culture. Survey data support this—a preference for legalizing marijuana has rapidly increased for all cohorts through *intracohort* change (Felson et al. 2019; cf. Kiley and Vaisey 2020:20).[21] The same pattern of change is true of the increasing acceptance of gay marriage (Rosenfeld 2017).[22] People's belief in the value of marriage may not change, even as their type of "marriage" might adapt to include same-sex romantic partnerships. These are two of the fastest changing attitudes that can be explained by a combination of settled dispositions and active updating, only what is *updated* is not attitudes per se, but ***types***.

Like finding Perrier, sushi, and crème fraîche in a Trump supporter's refrigerator, people with staunch anti-drug positions may confront friends who use marijuana to manage chemo, or know people who are not "criminals" who smoke pot recreationally, or learn about how marijuana laws have been unequally applied to people of color. The more they classify these people or behaviors as not criminal or laws as unfair, they may reclassify marijuana as no different from alcohol. Over time, repeated reclassification will likely lead people to reconfigure and revise their types.

Thinking in terms of interactions between attitudes and types offers a compromise position between the settled dispositions and active updating model. In this

[21] https://www.pewresearch.org/fact-tank/2016/10/12/support-for-marijuana-legalization-continues-to-rise/

[22] https://www.pewresearch.org/politics/2013/03/20/growing-support-for-gay-marriage-changed-minds-and-changing-demographics

way, we do not challenge or dismiss the wealth of evidence in support of the stability of attitudes, but *type reclassification and revision* offers a mechanism explaining changes in public opinion, policy, and law *despite* that stability. Taylor et al. (2019) argued that attention to type–token dynamics can reveal how types change *in the service of stability*. Sometimes to keep a disposition settled, reclassification and maintenance of types are necessary to bolster or stabilize dispositions, and at the same time they make possible cultural change at the personal (revision of type) and public (changes in public opinion) level. We call for a new focus on type–token revision and (re)classification, as essential to tracing cultural change by incorporating the stabilities and instabilities of environments. Through attention to how types expand, contract, reconfigure, and emphasize different qualities of tokens, we identify a mechanism for cultural change at the personal and public levels, even as attitudes may stay relatively constant.

This study is limited in that we have not assessed whether these reclassifications are sticky, or if personal culture is revised in the "long term." Future work on reclassification should assess the durability of type reclassifications and what conditions might increase durability. However, we want to stress that "real" change in personal culture is not only whether an individual is consistent across multiple waves of a study. Changing one's mind, even if short-term, may be socially significant with long-term consequences. If, for example, it occurs at key junctures in the political process (e.g., Brexit vote), stokes a general doubt in an institution (Michaels 2008), or leads one to question one's own sanity (Sweet 2019).

Implications for Persuasion

How can people leverage culture to persuade people to think and act for the common good? Cultural sociology often does not have optimistic answers to this question. While we know persuasion and personal transformation happens, it seems to happen rarely. Evidence against persuasion can be found in both the stability of dispositions and in the complexities of the environment through "cultural entropy" (McDonnell 2016) suggesting that persuasion projects face many obstacles to be successful. People who were dispositionally inclined to follow COVID mask and distancing mandates followed them, and those who were not inclined did not. This is a case of cultural entropy where the domain of politics complicated health mandates and super-charged the stakes of masks in ways experts did not predict. In a conversation with Anthony Fauci (director of the National Institute of Allergy and Infectious Disease), his boss Francis Collins (director of the National Institutes of Health) made this clear: "To see the way in which this whole set of public health recommendations took on a very strong political overtone that didn't belong there was incredibly disheartening ... the simple measures that really should have been adhered to were so difficult to get people to actually appreciate in the environment cast doubts on were picked up by people" (Kelly 2021).[23]

[23] https://www.npr.org/2021/03/09/975263049/in-the-24th-mile-of-a-marathon-fauci-and-collins-reflect-on-their-pandemic-year

These findings therefore have implications for our understanding of persuasion. Even if persuasion is rare, we should work to identify when and how it occurs. Shifting attitudes may be difficult but altering people's *classifications* may be less burdensome. Using a case of refrigerators, where the type is well established but the tokens are ambiguous offers a baseline for understanding this form of change. For a low-stakes classification task, we successfully encouraged some significant reclassification in the definitive condition. Higher stakes tasks, with more familiar tokens through routine or expertise might yield even less movement, suggesting just how difficult widespread persuasion really is, though with exposure to more tokens and definitive feedback we would likely see more reclassification until the type is revised and re-stabilized.

We have argued that reclassification is one form of cultural change at the level of personal culture. We have further argued that type–token relationships might be more moveable than more settled dispositions such as attitudes and values. As the Centers for Disease Control and Prevention (CDC) seek to persuade the public to view vaccines as "safe" rather than "risky," as police departments are encouraged to push back against implicit bias associating race and crime, as scientists try to convince the public that global warming is real, these findings offer practical guidance on language that might lead to reclassifying. Take climate science coverage as an example. We have a partial account for why doubts about global warming are perpetuated in the public sphere. Often the press uses language like "most climate scientists agree" or that "97% of Climate Scientists" believe in global warming making it easier for people to deny those claims and see a "debate" about "opinion" (see Cox 2013). Such normative language and "bothsidesism" may have opened up space for doubling down on climate denialism and embracing an embattled minority opinion. Had reporters used direct and definitive language from the start it may have limited these tendencies.

CONCLUSION

We have argued that one way that culture changes is through type–token dynamics: how qualities of tokens are aligned (or not) with types and the feedback people receive. These dynamics, in turn, lead to different processes of meaning making, like reclassifying a token from one type to another. This helps to clarify a puzzle at the center of work on cultural change. If dispositions appear stable, but cultural change happens, what is the mechanism when cohort replacement is an insufficient explanation? We offer type–token dynamics as a possible answer. Small changes in classification may reinforce dispositions while still yielding changes in public opinion and policy. Changes in the environment may rarely upend values or attitudes, but more regularly challenge people's typifications. These challenges then subtly reconfigure what counts as good or bad, criminal or innocent, scientific fact or "hooey," even as people's attitudes and commitments to the good, to justice, or to science remain unchanged. As such, we have proposed type–token dynamics as a mechanism and reclassification as a process that accounts for how dispositions appear stable over time even as cultural change happens that cannot be explained by cohort replacement. A new research agenda is needed to further these links between classification, cognition, culture, and change, one that tests the possibility of big effects from reclassification, type revision, and small changes in personal culture.

REFERENCES

Baumeister, Roy F., Ellen Bratslavsky, Catrin Finkenauer, and Kathleen D. Vohs. 2001. "Bad Is Stronger Than Good." *Review of General Psychology: Journal of Division 1, of the American Psychological Association* 5: 4: 323–370.

Bourdieu, Pierre. 1977. *Outline of a Theory of Practice*. Cambridge: Cambridge University Press.

Bourdieu, Pierre. 1984. *Distinction: A Social Critique of the Judgement of Taste*. London: Routledge & Kegan Paul.

Bourdieu, Pierre. 1988. *Homo Academicus*. Stanford, CA: Stanford University Press.

Buhrmester, Michael, Tracy Kwang, and Samuel D. Gosling. 2011. "Amazon's Mechanical Turk: A New Source of Inexpensive, Yet High-Quality, Data?" *Perspectives on Psychological Science: A Journal of the Association for Psychological Science* 6: 1: 3–5.

Casler, Krista, Lydia Bickel, and Elizabeth Hackett. 2013. "Separate but Equal? A Comparison of Participants and Data Gathered via Amazon's MTurk, Social Media, and Face-to-Face Behavioral Testing." *Computers in Human Behavior* 29: 6: 2156–2160.

Cerulo, Karen A. 2002. *Culture in Mind: Toward a Sociology of Culture and Cognition*. New York, NY: Routledge.

Cerulo, Karen A. 2006. *Never Saw It Coming: Cultural Challenges to Envisioning the Worst*. University of Chicago Press.

Cerulo, Karen A. 2010. "Mining the Intersections of Cognitive Sociology and Neuroscience." *Poetics* 38: 2: 115–132.

Cerulo, Karen A. 2018. "Scents and Sensibility: Olfaction, Sense-Making, and Meaning Attribution." *American Sociological Review* 83: 2: 361–389.

Chemero, Anthony. 2003. "An Outline of a Theory of Affordances." *Ecological Psychology: A Publication of the International Society for Ecological Psychology* 15: 2: 181–195.

Correll, Shelley J., Cecilia L. Ridgeway, Ezra W. Zuckerman, Sharon Jank, Sara Jordan-Bloch, and Sandra Nakagawa. 2017. "It's the Conventional Thought That Counts: How Third-Order Inference Produces Status Advantage." *American Sociological Review* 82: 2: 297–327.

Cox, J. Robert. 2013. "Media Convergence, Climate Science, and Public Uncertainty," In Libby Lester and Brett Hutchins (eds.), *Environmental Conflict and the Media*: pp. 231–243. New York, NY: Peter Lang.

DellaPosta, Daniel, Yongren Shi, and Michael Macy. 2015. "Why Do Liberals Drink Lattes?" *AJS; American Journal of Sociology* 120: 5: 1473–1511.

DiMaggio, Paul. 1997. "Culture and Cognition." *Annual Review of Sociology* 23: 1: 263–287.

Douglas, Mary. [1966] 2003. *Purity and Danger: An Analysis of Concepts of Pollution and Taboo*. Routledge.

Durkheim, Emile. [1912] 1995. *The Elementary Forms of the Religious Life* Translated by Karen Fields. New York, NY: Free Press.

Durkheim, Emile and Marcel Mauss. [1963] 2009. *Primitive Classification*. Routledge.

Eil, David and Justin M. Rao. 2011. "The Good News-Bad News Effect: Asymmetric Processing of Objective Information about Yourself." *American Economic Journal: Microeconomics* 3: 2: 114–138.

Engman, Athena and Cynthia Cranford. 2016. "Habit and the Body: Lessons for Social Theories of Habit From the Experiences of People With Physical Disabilities." *Sociological Theory* 34: 1: 27–44.

Felson, Jacob, Amy Adamczyk, and Christopher Thomas. 2019. "How and Why Have Attitudes about Cannabis Legalization Changed so Much?" *Social Science Research* 78: 12–27.

Firat, Rengin B., Hye Won Kwon, and Steven Hitlin. 2018. "A Novel Measure of Moral Boundaries: Testing Perceived In-group/Out-Group Value Differences in a Midwestern Sample." *Socius* 4: 2378023118818741.

FitzGerald, Chloë, Angela Martin, Delphine Berner, and Samia Hurst. 2019. "Interventions Designed to Reduce Implicit Prejudices and Implicit Stereotypes in Real World Contexts: A Systematic Review." *BMC Psychology* 7: 1: 29.

Friedman, Asia. 2011. "Toward a Sociology of Perception: Sight, Sex, and Gender." *Cultural Sociology* 5: 2: 187–206.

Garfinkel, Harold. 1967. *Studies in Ethnomethodology*. Englewood Cliffs, NJ: Prentice-Hall.

Gartman, David. 2007. "The Strength of Weak Programs in Cultural Sociology: A Critique of Alexander's Critique of Bourdieu." *Theory and Society* 36: 5: 381–413.
Gibson, James J. 2014. *The Ecological Approach to Visual Perception: Classic Edition*. Psychology Press.
Guilbeault, Douglas, Andrea Baronchelli, and Damon Centola. 2021. "Experimental Evidence for Scale-Induced Category Convergence Across Populations." *Nature Communications* 12: 1: 327.
Hattie, John and Helen Timperley. 2007. "The Power of Feedback." *Review of Educational Research* 77: 1: 81–112.
Hebart, Martin N., Charles Y. Zheng, Francisco Pereira, and Chris I. Baker. 2020. "Revealing the Multidimensional Mental Representations of Natural Objects Underlying Human Similarity Judgements." *Nature Human Behaviour* 4: 11: 1173–1185.
Hunzaker, M. B. Fallin. 2016. "Cultural Sentiments and Schema-Consistency Bias in Information Transmission." *American Sociological Review* 81: 6: 1223–1250.
Jacobs, Robert C. and Donald T. Campbell. 1961. "The Perpetuation of an Arbitrary Tradition Through Several Generations of a Laboratory Microculture." *Journal of Abnormal and Social Psychology* 62: 3: 649–658.
Kahn, Kimberly Barsamian and Paul G. Davies. 2011. "Differentially Dangerous? Phenotypic Racial Stereotypicality Increases Implicit Bias Among Ingroup and Outgroup Members." *Group Processes & Intergroup Relations: GPIR* 14: 4: 569–580.
Kahn, Kimberly Barsamian, Phillip Atiba Goff, J. Katherine Lee, and Diane Motamed. 2016. "Protecting Whiteness: White Phenotypic Racial Stereotypicality Reduces Police Use of Force." *Social Psychological and Personality Science* 7: 5: 403–411.
Kelly, Mary Louise. 2021 "In The '24th Mile' of A Marathon, Fauci And Collins Reflect on Their Pandemic Year." National Public Radio All Things Considered. March, 9. https://www.npr.org/2021/03/09/975263049/in-the-24th-mile-of-a-marathon-fauci-and-collins-reflect-on-their-pandemic-year
Kiley, Kevin and Stephen Vaisey. 2020. "Measuring Stability and Change in Personal Culture Using Panel Data." *American Sociological Review* 85: 3: 477–506.
Koffka, Kurt. 1935. *Principles of Gestalt Psychology*. New York: Harcourt, Brace.
Koriat, Asher, Morris Goldsmith, and Ainat Pansky. 2000. "Toward a Psychology of Memory Accuracy." *Annual Review of Psychology* 51: 481–537.
Lakoff, George. 2008. *Women, Fire, and Dangerous Things: What Categories Reveal About the Mind*. University of Chicago Press.
Lembo, Alessandra. 2020. "He Heard, She Heard: Toward a Cultural Sociology of the Senses." *Sociological Forum* 37: 115.
Leschziner, Vanina and Gordon Brett. 2019. "Beyond Two Minds: Cognitive, Embodied, and Evaluative Processes in Creativity." *Social Psychology Quarterly* 82: 4: 340–366.
Leschziner, Vanina and Adam Isaiah Green. 2013. "Thinking About Food and Sex Deliberate Cognition in the Routine Practices of a Field." *Sociological Theory* 31: 2: 116–144.
Lévi-Strauss, Claude. 1963. *Totemism*. Beacon Press.
Lizardo, Omar. 2004. "The Cognitive Origins of Bourdieu's Habitus." *Journal for the Theory of Social Behaviour* 34: 4: 375–401.
Lizardo, Omar. 2017. "Improving Cultural Analysis: Considering Personal Culture in Its Declarative and Nondeclarative Modes." *American Sociological Review* 82: 1: 88–115.
Lizardo, Omar, Robert Mowry, Brandon Sepulvado, Dustin S. Stoltz, Marshall A. Taylor, Justin Van Ness, and Michael Wood. 2016. "What Are Dual Process Models? Implications for Cultural Analysis in Sociology." *Sociological Theory* 34: 4: 287–310.
Lizardo, Omar and Michael Strand. 2010. "Skills, Toolkits, Contexts and Institutions: Clarifying the Relationship between Different Approaches to Cognition in Cultural Sociology." *Poetics* 38: 2: 205–228.
Luft, Aliza. 2020. "Theorizing Moral Cognition: Culture in Action, Situations, and Relationships." *Socius* 6: 2378023120916125.
Mandler, Jean Matter. 2004. *The Foundations of Mind: Origins of Conceptual Thought*. Oxford University Press.
Mandler, Jean M. 2008. "On the Birth and Growth of Concepts." *Philosophical Psychology* 21: 2: 207–230.
Margolis, Eric. 1994. "A Reassessment of the Shift from the Classical Theory of Concepts to Prototype Theory." *Cognition* 51: 1: 73–89.
Margolis, Eric and Stephen Laurence. 1999. *Concepts: Core Readings*. MIT Press.
Martin, John Levi. 2010. "Life's a Beach but You're an Ant, and Other Unwelcome News for the Sociology of Culture." *Poetics* 38: 2: 229–244.
Martin, John Levi. 2011. *The Explanation of Social Action*. USA: Oxford University Press.
Martin, John Levi. 2015. *Thinking Through Theory*. W.W. Norton.

McDonnell, Terence E. 2016. *Best Laid Plans: Cultural Entropy and the Unraveling of AIDS Media Campaigns*. University of Chicago Press.

Melamed, David, Leanne Barry, Bradley Montgomery, and Oneya F. Okuwobi. 2020. "Measuring Racial Status Beliefs with Implicit Associations." *American Sociological Review* 85: 6: 1123–1131.

Metcalfe, Janet. 2017. "Learning From Errors." *Annual Review of Psychology* 68: 465–489.

Michaels, David. 2008. *Doubt Is Their Product: How Industry's Assault on Science Threatens Your Health*. Oxford University Press.

Morris, Michael W. and Paul C. Moore. 2000. "The Lessons We (Don't) Learn: Counterfactual Thinking and Organizational Accountability after a Close Call." *Administrative Science Quarterly* 45: 4: 737–765.

Nolan, Daniel A. 2020. "Souvenirs and Travel Guides: Cognition, Culture, and Grieving Public Figures." *Poetics* 84: 101484.

Patterson, Orlando. 2004. "Culture and Continuity: Causal Structures in Socio-Cultural Persistence." *Matters of Culture: Cultural Sociology in Practice* 40: 71–109.

Patterson, Orlando. 2013. "Making Sense of Culture." *Annual Review of Sociology* 40: 1: 140512172738005.

Peeters, Guido. 1971. "The Positive-Negative Asymmetry: On Cognitive Consistency and Positivity Bias." *European Journal of Social Psychology* 1: 4: 455–474.

Ridgeway, Cecilia L. and Shelley J. Correll. 2006. "Consensus and the Creation of Status Beliefs." *Social Forces; a Scientific Medium of Social Study and Interpretation* 85: 1: 431–453.

Rosch, Eleanor H. 1973. "Natural Categories." *Cognitive Psychology* 4: 3: 328–350.

Rosch, Eleanor. 1999. "Principles of Categorization". *Concepts: Core Readings* 189: 312–322.

Rosch, Eleanor and Carolyn B. Mervis. 1975. "Family Resemblances: Studies in the Internal Structure of Categories." *Cognitive Psychology* 7: 4: 573–605.

Rose-Greenland, Fiona. 2016. "Color Perception in Sociology Materiality and Authenticity at the Gods in Color Show." *Sociological Theory* 34: 2: 81–105.

Rosenfeld, Michael J. 2017. "Moving a Mountain: The Extraordinary Trajectory of Same-Sex Marriage Approval in the United States." *Socius* 3: 2378023117727658.

Sahlins, Marshall. 1976. *Culture and Practical Reason*. University of Chicago Press.

Sedikides, Constantine and John J. Skowronski. 2020. "In Human Memory, Good Can Be Stronger Than Bad." *Current Directions in Psychological Science* 29: 1: 86–91.

Shepherd, Hana. 2010. "Classification, Cognition and Context: The Case of the World Bank." *Poetics* 38: 2: 134–150.

Sherif, Muzafer. 1935. "A Study of Some Social Factors in Perception." *Archives of Psychology (Columbia University)* 187: 60.

Smith, Robert J. 2015. "Reducing Racially Disparate Policing Outcomes: Is Implicit Bias Training the Answer." *University of Hawai'i Law Review* 37: 2: 295.

Sobotka, Tagart Cain. 2020. "Not Your Average Joe: Pluralistic Ignorance, Status, and Modern Sexism." *Men and Masculinities* 1097184X20901578.

Stoltz, Dustin S. and Marshall A. Taylor. 2017. "Paying with Change: The Purposeful Enunciation of Material Culture." *Poetics* 64: 26–39.

Strand, Michael and Omar Lizardo. 2017. "The Hysteresis Effect: Theorizing Mismatch in Action." *Journal for the Theory of Social Behaviour* 47: 2: 164–194.

Strauss, Claudia and Naomi Quinn. 1997. *A Cognitive Theory of Cultural Meaning*, vol 9. Cambridge University Press.

Sutton, John, Celia B. Harris, Paul G. Keil, and Amanda J. Barnier. 2010. "The Psychology of Memory, Extended Cognition, and Socially Distributed Remembering." *Phenomenology and the Cognitive Sciences* 9: 4: 521–560.

Sweet, Paige L. 2019. "The Sociology of Gaslighting." *American Sociological Review* 84: 5: 851–875.

Taylor, John. 2003. "The Classical Approach to Categorization," In John Taylor (ed.), *Linguistic Categorization*: pp. 19–40. Oxford: Oxford University Press.

Taylor, Marshall A., Dustin S. Stoltz, and Terence E. McDonnell. 2019. "Binding Significance to Form: Cultural Objects, Neural Binding, and Cultural Change." *Poetics* 73: 1–16.

Taylor, Shelley E. 1991. "Asymmetrical Effects of Positive and Negative Events: The Mobilization-Minimization Hypothesis." *Psychological Bulletin* 110: 1: 67–85.

Turner, Victor. 1970. *The Forest of Symbols: Aspects of Ndembu Ritual*. Cornell University Press.

Vaisey, Stephen. 2009. "Motivation and Justification: A Dual-Process Model of Culture in Action 1." *The American Journal of Sociology* 114: 6: 1675–1715.

Vaisey, Stephen and Kevin Kiley 2021. "A Model-Based Method for Detecting Persistent Cultural Change Using Panel Data." *Sociological Science* 8: 83–95.

Vaisey, Stephen and Omar Lizardo. 2016. "Cultural Fragmentation or Acquired Dispositions? A New Approach to Accounting for Patterns of Cultural Change." *Socius* 2: 2378023116669726.

Williams, Richard. 2012. "Using the Margins Command to Estimate and Interpret Adjusted Predictions and Marginal Effects." *The Stata Journal* 12: 2: 308–331.

Wood, Michael Lee, Dustin S. Stoltz, Justin Van Ness, and Marshall A. Taylor. 2018. "Schemas and Frames." *Sociological Theory* 36: 3: 244–261.

Zadeh, Lofti Asker. 1996. "A Note on Prototype Theory and Fuzzy Sets," In George J. Klir and Bo Yuan (eds.), *Fuzzy Sets, Fuzzy Logic, and Fuzzy Systems*: Vol. 6, pp. 587–593, *Advances in Fuzzy Systems? Applications and Theory*. Singapore: World Scientific.

Zerubavel, Eviatar 1996. "Lumping and Splitting: Notes on Social Classification." *Sociological Forum* 11: 3: 421–433.

Zucker, Lynne G. 1977. "The Role of Institutionalization in Cultural Persistence." *American Sociological Review* 42: 5: 726–743.