Social categories create (biased) semantic interference during face naming

Dominic J. Packer (djp208@lehigh.edu)
Psychology and Cognitive Science, Lehigh University, 17 Memorial Dr. E.
Bethlehem, PA 18015 USA

Pádraig G. O'Séaghdha (pat.oseaghdha@lehigh.edu)
Psychology and Cognitive Science, Lehigh University, 17 Memorial Dr. E.
Bethlehem, PA 18015 USA

Almut Hupbach (hupbach@lehigh.edu)
Psychology and Cognitive Science, Lehigh University, 17 Memorial Dr. E.
Bethlehem, PA 18015 USA

Joseph E. Bates (x5a512@lehigh.edu)
Psychology, Lehigh University, 17 Memorial Dr. E.
Bethlehem, PA 18015 USA

Abstract

Semantic interference in word retrieval has been observed for both well-learned and ad hoc inter-item relations. We tested whether such semantic interference extends to the blocked cyclic naming of racially homogeneous vs. heterogeneous faces. No information except arbitrarily assigned names was provided for novel faces. Yet we observed interference in naming individuals in homogeneous groups. Moreover, consistent with other findings in the social domain, interference occurred for other-race but not for own-race faces. Because this interference effect does not require a rich knowledge base about individuals, it is consistent with the view that interference arises in adjustments to the strength of conceptual-lexical links rather than in knowledge structures themselves. Evidence of modulation by target race further suggests that interference effects may provide an effective tool for exploration of social categorization processes.

Keywords: Semantic interference; language production; blocked cyclic naming; social categorization; intergroup bias

Introduction

This paper presents an initial investigation of how basic memory retrieval and language production processes are affected by social context. When people name objects, they often exhibit semantic interference in which retrieving a target word from memory disrupts retrieval of words that belong to the same semantic category (e.g., Damian, Viglocco & Levelt, 2001). In the current research, we examined whether similar interference effects occur when naming faces of members from social categories, specifically, racial groups. We further investigated whether people exhibit a 'name retrieval bias', such that there is greater interference when naming other-race compared to own-race faces.

Semantic Interference in Language Production

Semantic interference highlights the competitive nature of word selection in language production. In the blocked-cyclic naming paradigm, for example, participants repeatedly name small sets of pictures (e.g., four pictures each named individually four times). The pictures are either presented in homogenous blocks in which they share a common semantic relation, or in heterogeneous blocks in which they do not have identifiable semantic links. Naming times are slower in homogeneous than heterogeneous blocks, and the level of interference often increases over cycles (e.g., Damian et al., 2001; Schnur et al., 2009). Semantic interference occurs in this and similar paradigms because retrieving a word co-activates semantically related words, which compete with and slow selection of the target word. Further, retrieving a word primes its subsequent retrieval, making it a stronger competitor when later naming related words (Howard, Nickels, Coltheart & Cole-Virtue, 2006). The change in lexical accessibility is long lasting, suggesting that a learning mechanism that damps accessibility of competitors while strengthening the current item, rather than short-term modulation of activation, is at the core of semantic interference (Oppenheim, Dell & Schwartz, 2010; see Navarrete, DelPrato & Mahon, 2012). fMRI studies have localized lexical selection to left inferior frontal gyrus [LIFG] (believed to be involved in competition resolution) and linked areas of temporal cortex (Schnur et al., 2009).

Traditionally, semantic interference experiments have investigated how shared membership in fixed taxonomic categories (e.g., animals, vegetables, minerals) generates interference. In these cases, items share both category memberships and overlapping semantic features (e.g., legs, heads, locomotion). However, recent research has also shown interference for items linked by a semantic theme

(e.g., garden links slug, gardener and rake; see Abdel Rahman & Melinger, 2007), as well as for items that are linked as members of an ad hoc category. Abdel Rahman and Melinger (2011) had participants complete a cyclic-blocked naming task with pictures that had no obvious semantic relation to one another (e.g., stool, knife, bucket, river), but that could be combined as members of an ad hoc category (e.g., "things present on a fishing trip"). No interference was found when participants were unaware of the ad hoc category; however, interference (longer naming latencies) arose when participants were informed about the category. These findings demonstrate the highly dynamic nature of the semantic system, such that items without fixed shared semantic features nevertheless rapidly become associated (i.e., exhibit shared activation and competition) when linked by a thematic context.

Effects of Social Categories

Social categories (e.g., racial group memberships) function in many ways like other categories. People tend, for example, to exaggerate within group similarities and between group differences for both social and non-social categories (e.g., Levin & Angelone, 2002; Tajfel & Wilkes, 1963). As such, by virtue of their category membership, individuals are assumed to possess common features, and are often stereotyped accordingly (e.g., Kunda & Spencer, 2003). Further, the effects of social categories often emerge rapidly and automatically (e.g., Devine, 1989). For example, ERP studies have observed category-based differences in neural signals associated with early visual processing of different race faces (Ito & Bartholow, 2009). In these studies, white participants show heightened P100 and N170 responses (which have been linked to early face processing) when viewing White versus Black faces (e.g., Ito & Urland, 2003; Cunningham, Van Bavel, Arbuckle, Packer & Waggoner, 2012). The robust influence of social categories extends to a wide range of cognitive and affective processes, ultimately shaping behaviors including affiliation, cooperation and conflict (Turner, Hogg, Oakes, Reicher & Wetherell, 1987).

Prior research has also shown effects of social categories on language use, such that people strategically use language to enhance ingroups and derogate outgroups. In a phenomenon known as the 'linguistic intergroup bias', people tend to describe positive ingroup and negative outgroup behaviors more abstractly than they describe negative ingroup and positive outgroup behaviors (Maass, 1999; Maass, Salvi, Arcuri & Semin, 1989). The use of relatively abstract words (e.g., adjectives – helpful, aggressive) to communicate "our" desirable and "their" undesirable actions implies that these are enduring and global characteristics. In contrast, using relatively concrete words (e.g., action verbs – help, hit) to communicate our undesirable and their desirable actions conveys that these behaviors are situationally-specific and transient.

In the current research, we investigated whether social categories affect basic language production processes under controlled experimental conditions. We did this in the domain of face naming. Specifically, we asked whether shared social categories induce semantic interference effects during person (face) naming. We had participants learn the names of 16 novel faces belonging to four different racial groups (four faces in each). They then completed a blocked-cyclic naming task with these faces. In homogenous blocks, participants cycled through naming four faces that all belonged to the same racial group; in heterogeneous blocks, participants cycled through naming four faces that each belonged to a different racial group.

In contrast to common objects, faces are processed through partly specialized cortical networks, including fusiform gyrus. In addition, because person names are arbitrary, their retrieval from face configurations may be more difficult than object naming (e.g., Valentine, Brennen & Brédart, 1996; see also Griffin, 2010). However, given that social categories exert robust effects on a wide range of psychological processes and function similarly to non-social categories, we expected to observe semantic interference in face naming. Due to their common category membership (and overlapping visual and possibly semantic features – e.g., stereotypes), retrieving the name of one group member should increase co-activation of the names of other members, which will compete with and slow selection of the target name.

Social Categorization Biases

Social categories are not entirely analogous to non-social categories. In particular, people often exhibit intergroup biases, such that members of ingroups and outgroups are processed differently (e.g., the linguistic intergroup bias described above). These biases come in different forms, but many are reducible to the observation that outgroup members tend to be processed more categorically than ingroup members. Whereas ingroup members are typically individuated and treated as distinct entities, outgroup members are often treated as relatively interchangeable exemplars of their group (Brewer, 1988; Fiske & Neuberg, 1990).

One such bias emerges in facial recognition. The 'other race effect' (or 'own race bias') refers to the well-replicated finding that people are generally better at recognizing members of their own versus other racial groups (e.g., in incidental recognition paradigms). Although perceptual expertise is a contributor (i.e., people typically have more experience processing own than other race faces), recent research suggests that this bias is largely a categorization driven effect (Hugenberg, Young, Bernstein & Sacco, 2010; Van Bavel, Packer & Cunningham, 2011). According to Hugenberg et al.'s (2010) Categorization-Individuation Model, for example, classifying faces as exemplars of a category focuses attention on category-diagnostic (shared) features, which reduces subsequent ability to discriminate among category members. In contrast, when faces are not categorized but instead individuated, attention is focused on distinct features, enhancing the ability to discriminate them

later. Critically, intergroup biases in facial recognition emerge because outgroup faces tend to activate their categories more strongly than ingroup faces (Hugenberg et al., 2010; Levin, 1996, 2000; Stroessner, 1996).

We anticipated that a similar bias might also occur in name production. If outgroup faces invoke categorization more strongly than ingroup faces, semantic interference should be stronger for outgroup than ingroup faces. Retrieving an outgroup member's name should co-activate the names of other outgroup members. In contrast, because ingroup members tend to be individuated, retrieving the name of an ingroup member should not increase activation of other ingroup members' names, at least not to the same extent. If so, homogenous ingroup naming may not differ from heterogeneous condition naming.

To summarize, we predicted semantic interference during naming of faces when they are racially grouped. Based on ingroup/outgroup differences in social-cognitive processing, we further hypothesized that such interference would be stronger for other-race than own-race faces.

Methods

Participants

Eighteen introductory psychology students at Lehigh University participated for partial course credit. This sample size provides good power because the repeated measures design collects many observations from each participant. The average age was 19.22 years, and there were 9 males and 8 females (one did not report gender). All participants spoke English as a first language, and reported European ethnic origins during pre-testing. A racially homogeneous sample was important in this case for testing hypothesized ingroup/outgroup differences. During testing, one participant indicated a mixed ethnic background (European and Asian). Analyses including vs. excluding this participant yielded identical findings; we therefore report analyses including all participants.

Design

We used a 2 (Context: heterogeneous, homogeneous) X 4 (Race: Asian, Black, Middle-Eastern, White) X 4 Replication (1, 2, 3, 4) within subjects design.

Procedure

Learning Phase. After familiarization with the picture naming set-up, participants first learned arbitrarily assigned names for 16 male faces. Four faces belonged to each of four racial groups: Asian, Black, Middle Eastern and White. All names were single syllable, of European origin, and common in the North American context. Each name also started with a different letter (e.g., Bill, Chris, Dan). The names were assigned to faces in two different randomizations, which were counter-balanced across participants. Participants initially viewed the 16 face/name pairings in a randomized 4 X 4 matrix on the computer screen for 2 minutes, and were instructed to try to memorize as many as they could. Each face/name was then presented twice for two seconds in random order, with participants instructed to read the names aloud.

Testing Phase. The testing phase consisted of four replications, each containing eight sets of 16 trials. Each set comprised four faces repeated semi-randomly (for each participant) across four repetition cycles. Four of the sets were racially homogenous, four were racially heterogeneous, and the order of sets within each replication was randomly determined for each participant. In total, participants completed 512 trials.

At the beginning of each set, participants were shown four faces along with their names and were asked to read each name aloud. Then, to confirm that they remembered the names, they were presented with each face individually (in random order) and were asked to provide the name (which appeared on the screen upon vocalization to confirm or correct participants' responses). This was repeated until participants named all four faces correctly. In most cases, no repetitions were required.

Each trial began with a fixation cue (*) displayed for 100 -milliseconds (ms), along with a warning sound, followed by a face. Naming latencies were measured with a voice key. The face remained on the screen until a name was produced, or for a maximum of 1500 ms. After naming, the face disappeared and was followed by a blank screen for 1500 ms. Participants were instructed to speak clearly, and to name the faces as quickly and accurately as possible.

Results

Following standard practice for this type of design, the heterogeneous context responses in each replication were sorted to match the corresponding homogeneous groupings. We then conducted a 2 (Context: heterogeneous, homogeneous) X 4 (Race: Asian, Black, Middle-Eastern, White) X 4 Replication (1, 2, 3, 4) analysis on the speed with which participants named faces. Specifically, we implemented a multi-level model in which trials were nested within participants using the PROC MIXED procedure in SAS. Multi-level models allow for more accurate estimates of effects by accounting for interdependence among trials within participants. We removed error trials on which participants named a face incorrectly (1.7%), the voice key was triggered by something other than a name (e.g., a cough or stutter, 0.9%) or participants did not respond within the time window (1.0%). We also removed trials with RTs < 200ms.

Extending prior research on semantic interference effects in blocked cyclic naming paradigms, there was a significant main effect of Context, $F(1, 17) = 20.05$, $p < .001$. Overall, participants were slower to name faces in racially homogeneous ($M = 701$, $SD = 179$) versus heterogeneous contexts ($M = 671$, $SD = 160$). Critically, however, the effect of Context was modulated by Race [Context x Race interaction: $F(3, 51) = 3.56$, $p < .05$]. As shown in Figure 1A, naming faces in homogenous (vs. heterogeneous)

contexts produced statistically significant interference effects for Asian ($p < .05$), Black and Middle-Eastern faces ($ps < .01$). However, there was no evidence of an effect for White faces ($p > .30$). The Context X Race interaction was not moderated by Replication ($F(9, 107) = 0.85$, $p > .50$), indicating that the pattern was stable throughout the experiment.

Although they were relatively rare, examination of naming errors showed that they exhibited the same pattern as the reaction time data. Specifically, naming errors were more frequent in homogeneous than heterogeneous contexts for faces of all races except White (see Figure 1B)

Discussion

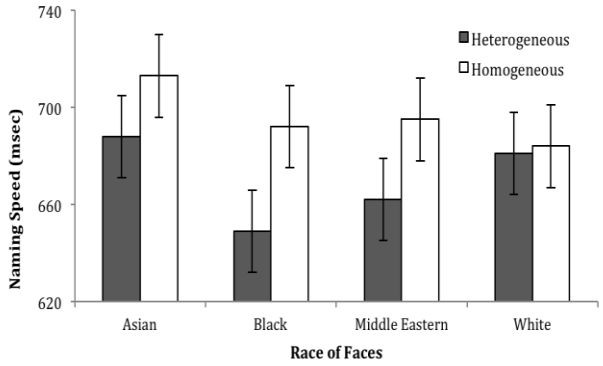
Our study shows several novel findings. First, we observed semantic interference in proper name retrieval in blocked cyclic naming. To our knowledge, this has not been reported previously. Second, the basis of the interference was racial grouping of the faces, extending previous reports of semantic interference among taxonomic, thematic or ad hoc associates to social categories instantiated by facial features. In this domain, the basis for interference in name retrieval is quite slender, comprising modulation of the mapping from face to name by the mere knowledge that the named individuals belong to a distinct racial group. Third, the effect was present for three "other race" groups, but was absent for the "own race" of the white participants. This is interesting both as a new manifestation of the own race bias (i.e., a name retrieval bias), and as evidence that semantic interference does not arise under all conditions. The set of white faces could certainly be construed as a category in the context of this experiment, and yet we observed no evidence of interference. The error data even suggest that naming of homogeneous white faces may have been facilitated.

Previous researchers of semantic interference have taken pains to show that the effect is not an artifact of visual similarity (e.g., Damian et al., 2001). This concern also arises in the case of face naming, because racially homogeneous faces might be more difficult to discriminate. However, the data do not support this possibility. The visual similarity explanation would predict a context effect for all of the groups (not the case), and greater difficulty in naming outgroup than ingroup members in heterogeneous contexts (also not the case). Thus our findings are clearly driven by categorization of outgroup faces and not by a perceptual similarity confound.

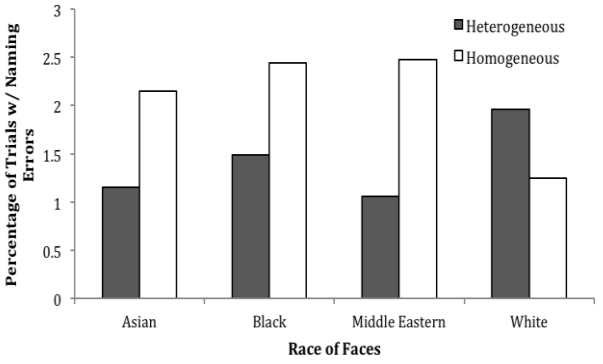
One way to interpret our findings (and link them to previous findings with ad hoc categories) is that an autonomous face recognition process is followed by categorically constrained name retrieval. Against this, however, is the finding of Van Bavel et al. (2011) that social categories other than race affect the functioning of the fusiform face area and that nonracial group affiliations can even trump visually salient characteristics such as race. If context dynamically modulates the functioning of face processing areas, it may be more appropriate to conceive of

Figure 1: Response Times and Error Rates as a Function of Race and Heterogeneous or Homogeneous Context. (RTs and their standard errors (pooled) are estimated from the multi-level model).

A. Response Times



B. Face Naming Errors



semantic interference in name retrieval as engaging relatively extensive neural networks.

To our knowledge, the present study is the first to demonstrate interference in name retrieval for faces from different racial categories. Research currently under way in our research group attempts to shed light on the basis of this effect by teasing apart whether common category membership, overlapping semantic features, or even shared visual features contribute to the effect. Another question we may examine in future is the cultural domain of the names. In this experiment, the names were all European in origin and they may have been perceived as more congruent with the White category. It is not clear, however, that this could account for the observed pattern of effects because tighter linkages among the names within the White category should presumably tend to increase rather than decrease interference and vice-versa for the other ethnic groups.

Additionally, we are investigating whether the observed name retrieval bias is specific to racial groups or extends to other social categories. Ongoing research is, for example,

using a minimal group paradigm in which participants are randomly assigned to novel and arbitrary groups (e.g., teams). Minimal groups trigger many of the same biases as other social categories (e.g., in face recognition; Van Bavel et al., 2011), and we anticipate that they may in this domain as well. Importantly, minimal groups do not differ systematically in visual features (e.g., members of all groups can belong to the same race), and participants do not possess semantic information (e.g., stereotypes) about the groups. To the extent that similar effects are observed with minimal groups, it will illuminate the role that categorization per se can play in interference effects. Our research contributes to the mounting evidence that influences of categorization on interference effects in word/name retrieval are dynamic, shifting as a function of currently available or salient categories (Abdel Rahman & Melinger, 2011).

Wider Implications

The own race bias in facial recognition has profound and disturbing social implications. For example, difficulties distinguishing between members of other races may be a significant cause of eyewitness misidentification and wrongful conviction in criminal cases (Scheck, Neufeld & Dwyer, 2003). The current research suggests that a similar bias may occur in name retrieval, such that people have greater difficulty retrieving the names of other race individuals. Most of the time, the consequences of a name retrieval bias may be minor, but in certain contexts this bias could have pernicious effects. For example, teachers may be less likely to call on other race students, perhaps particularly if those students tend to be encountered proximally (e.g., seated together). The correlates and consequences of the name retrieval bias shown here merit further investigation.

Acknowledgments

This research was supported by a Lehigh University CORE (Collaborative Research Opportunity) award. We thank Nevada Heft for contributions to the project.

References

Abdel Rahman, R. & Melinger, A. (2007). When bees hamper the production of honey: Lexical interference from associates in speech production. *Journal of Experimental Psychology: Learning, Memory, and Cognition, 33*, 604-614.

Abdel Rahman, R. & Melinger, A. (2011). The dynamic microstructure of speech production: Semantic interference built on the fly. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, *37*, 149-161.

Brewer, M. B. (1988). A dual process model of impression formation. In T. Srull & R. Wyer (Eds.), *Advances in Social Cognition, Vol. 1* (pp. 1–36). Hillsdale, NJ: Erlbaum.

Cunningham, W. A., Van Bavel, J. J., Arbuckle, N. L., Packer, D. J., & Waggoner, A. S. (2012). Rapid social perception is flexible: approach and avoidance motivational states shape P100 responses to other-race faces. *Frontiers in Human Neuroscience, 6*.

Damian, M. F., Vigliocco, G., & Levelt, W. J. (2001). Effects of semantic context in the naming of pictures and words. *Cognition*, *81*, B77-B86.

Devine, P. G. (1989). Stereotypes and prejudice: Their automatic and controlled components. *Journal of Personality and Social Psychology*, *56*, 5-18.

Fiske, S. T., & Neuberg, S. L. (1990). A continuum of impression-formation, from category-base to individuating processes: Influence of information and motivation on attention and interpretation. *Advances in Experimental Social Psychology, 23*, 1-74.

Griffin, Z. M. (2010). Retrieving personal names, referring expressions, and terms of address. *Psychology of Learning and Motivation*, *53*, 345-387.

Howard, D., Nickels, L., Coltheart, M., & Cole-Virtue, J. (2006). Cumulative semantic inhibition in picture naming: Experimental and computational studies. *Cognition, 100*, 464-482.

Hugenberg, K., Young, S. G., Bernstein, M. J., & Sacco, D. F. (2010). The categorization-individuation model: An integrative account of the other-race recognition deficit. *Psychological Review*, *117*, 1168-1187.

Ito, T. A., & Bartholow, B. D. (2009). The neural correlates of race. *Trends in cognitive sciences*, *13*(12), 524-531.

Ito, T. A., & Urland, G. R. (2003). Race and gender on the brain: electrocortical measures of attention to the race and gender of multiply categorizable individuals. *Journal of Personality and Social Psychology*, *85*, 616-626.

Kunda, Z., & Spencer, S. J. (2003). When do stereotypes come to mind and when do they color judgment? A goal-based theoretical framework for stereotype activation and application. *Psychological Bulletin*, *129*, 522-544.

Levin, D. T. (1996). Classifying faces by race: The structure of face categories. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, *22*, 1364-1382.

Levin, D. T. (2000). Race as a visual feature: Using visual search and perceptual discrimination tasks to understand face categories and the cross-race recognition deficit. *Journal of Experimental Psychology: General*, *129*, 559-574.

Levin, D. T., & Angelone, B. L. (2002). Categorical perception of race. *Perception, 31*, 567-578.

Maass, A. (1999). Linguistic intergroup bias: Stereotype perpetuation through language. *Advances in Experimental Social Psychology*, *31*, 79-121.

Maass, A., Salvi, D., Arcuri, L., & Semin, G. R. (1989). Language use in intergroup contexts: The linguistic intergroup bias. *Journal of Personality and Social Psychology*, *57*, 981-993.

Navarrete, E., Del Prato, P., and Mahon, B. (2012).Factors determining semantic facilitation and interference in the cyclic naming paradigm. *Frontiers in Psychology*, *3,* Article 38, 1-15.

Oppenheim, G. M., Dell, G. S., & Schwartz, M. F. (2010). The dark side of incremental learning: A model of cumulative semantic interference during lexical access in speech production. *Cognition*, *114*, 227-252.

Scheck, B., Neufeld, P., & Dwyer, J. (2003). *Actual innocence: When justice goes wrong and how to make it right*. New American Library.

Schnur, T., Schwartz, M., Kimberg, D., Hirshorn, E., Coslett, H., & Thompson-Schill, S. (2009). Localizing interference during naming: Convergent neuroimaging and neuropsychological evidence for the function of Broca's area. *PNAS,* 106, 322-327.

Stroessner, S. J. (1996). Social categorization by race or sex: Effects of perceived non-normalcy on response times. *Social Cognition*, *14*, 247-276.

Tajfel, H., & Wilkes, A. L. (1963). Classification and quantitative judgement. *British Journal of Psychology*, *54*, 101-114.

Turner, J. C., Hogg, M. A., Oakes, P. J., Reicher, S. D., & Wetherell, M. S. (1987). *Rediscovering the Social Group: A Self-Categorization Theory.* Cambridge, MA: Blackwell.

Valentine, T., Brennen, T., & Brédart, S. (1996). *The cognitive psychology of proper names: On the importance of being Ernest.* London: Routledge.

Van Bavel, J. J., Packer, D. J., & Cunningham, W. A. (2011). Modulation of the fusiform face area following minimal exposure to motivationally relevant faces: evidence of in-group enhancement (not out-group disregard). *Journal of Cognitive Neuroscience*, *23*, 3343-3354.