The labelled container: Conceptual development of social category representations

Rebekah A. Gelpi[1], Suraiya Allidina[1], Daniel Hoyer[2,3],
and William A. Cunningham[1]
[1] Department of Psychology, University of Toronto
[2] Evolution Institute
[3] George Brown College

Pietraszewski contends that group representations that rely on a "containment metaphor" fail to adequately capture phenomena of group dynamics such as shifts in allegiances. We argue, in contrast, that social categories allow for computationally efficient, richly structured, and flexible group representations that explain some of the most intriguing aspects of social group behaviour.

Keywords: social categories, groups, inductive reasoning, social identity

Pietraszewski (this issue) offers a bottom-up approach to understanding the problem of group representation, describing how these representations could be constructed *ex nihilo* in the context of conflict by identifying patterns of interactions within triads. While this account provides a straightforward computational account of mental representations of intergroup conflict, it neglects important top-down influences such as induction and generalization. We suggest that these social group categorization processes may better account for core features of modern human group living, such as flexible, dynamic social identities and generalized trust in strangers in massive environments, than an approach that relies only on constructing group representations out of an event framework.

As bottom-up processing alone is unlikely to be realizable at scale given humans' limited cognitive resources, top-down processes likely must play some role in developing group representations. One such process is inductive reasoning, which allows us to efficiently learn complex concept and category knowledge from relatively sparse data (Kemp, Perfors, & Tenenbaum, 2007; Tenenbaum, Kemp, Griffiths, & Goodman, 2011). While categories can be quickly built from experience, we also receive rich, highly structured information about these categories in the form of generic language and category labels, quickly scaffolding our learning about abstract category structures, both non-social (Butler & Markman, 2014; Taverna, Padilla, Baiocchi, & Peralta, 2021) and social (Baron, Dunham, Banaji, & Carey, 2014; Gelman, Ware, & Kleinberg, 2010; Roberts, Ho, & Gelman, 2017). By representing groups as abstract, symbolic categories that capture statistical regularities and make probabilistic predictions about how members are likely to appear and act, we gain several insights that are not captured by a model which constructs group representations out of event frameworks alone. Relying on category and concept learning not only allows us to use a highly tractable, domain-general strategy to learn about and represent complex social groups, it also readily expands the predictive capacity of group assignment beyond conflict and reciprocity, to resolving highly abstract cooperation and coordination problems that complex societies must solve.

Beyond simply giving "rules for assignment" based on statistical regularities in shared features and behaviours, categories allow us to make predictions about how category members are likely to behave, and even to develop prescriptive norms about how they ought to behave (Roberts, Gelman, & Ho, 2017; Foster-Hanson & Rhodes, 2019), including group members' moral obligations to the group (Chalik & Rhodes, 2018, 2020). In the context of a conflict, knowledge about agents' social categories can allow observers to make inferences about what kinds of behaviours or group-constitutive roles are more or less likely for an agent to take; rather than (or in addition to) the roles determining the inferences about groups, the prior expectations around group memberships determine the inference of roles.

Indeed, these processes may be critical in the formation and representation of groups where most members are strangers, as members of modern social groups such as cities and countries are able to recognize only a small proportion of the group's total population with fine individual detail (Dunbar, 2010), and must regularly interact with group members about which one has no information beyond knowledge of perceptually or contextually inferred social categories (such as race, gender, or nationality). Relying on group-based norms and expectations allows us to make broad inferences about strangers that facilitate social prediction without need-

ing to represent these individuals or interactions in a deep way. The cultural transmission of norms regarding social roles and division of labour (e.g., Lew-Levy, Lavi, Reckin, Cristóbal-Azkarate, & Ellis-Davies, 2018) can also easily be accommodated within a framework of social category learning and, we argue, is in fact essential to explain how individuals access and internalize these norms within large, complex societies (Gavrilets & Richerson, 2017).

In contrast to an event model, categorization processes are likely more involved in the social learning processes that underlie cultural learning. Rather than being learned from the bottom up, many group identities have been developed over long histories of cooperation, affiliation, or conflict. Although processes of social categorization are efficient for making sense of the social world, they come at a cost; when applied to these "prepackaged" groups, the cognitive tools of induction, categorization, and pedagogical reasoning that allow us to swiftly bootstrap sophisticated social category structures can lead to overgeneralization and stereotyping (Macrae, Milne, & Bodenhausen, 1994).

Interestingly, despite Pietraszewski's claim that categorizations are rigid, our social categories are not only swiftly and efficiently learned, but also highly flexible, dynamic, and complex in structure, with people holding multiple social identities which can overlap to varying degrees or even be contradictory (Roccas & Brewer, 2002). Social identities can rapidly shift in accordance with their context or prominence, and these shifts can in turn shape attention, modulating our memory for social categories (Van Bavel, Packer, & Cunningham, 2011; Van Bavel & Cunningham, 2012) as well as amygdala activation (Van Bavel, Packer, & Cunningham, 2008). When group identities are multiply nested or produce conflicting loyalties, contextual activation or prominence of a group identity could lead to using different social category structures to make behavioural decisions as well as predictions about likely outcomes in complex, real-world contexts.

We applaud Pietraszewski's goal of formalizing the process of social group representation and grounding our understanding of groups in this psychological process. While we agree that triadic event information can motivate people's reasoning about some—particularly ad hoc—groupings, we believe that representing groups as categories provides unique insight into phenomena such as the transmission of cultural knowledge and generalized trust while allowing for symbolic, abstract group identities that are nevertheless flexible and situationally dependent. These features, we argue, may help us better understand the development of "prepackaged" group identities such as nationalities and ethnic groups through our evolutionary history as well as represent how specific identities emerge and shift among contemporary populations. We propose that extending existing computational models of categorization and concept learning (e.g. Love, Medin, & Gureckis, 2004) into the social domain by

including dimensions such as identity or motivation may offer a highly tractable account for how we can deploy basic, domain-general inference processes to solve the challenges of reasoning about large and complex groups.

References

Baron, A. S., Dunham, Y., Banaji, M., & Carey, S. (2014). Constraints on the Acquisition of Social Category Concepts. *Journal of Cognition and Development*, *15*(2), 238–268. doi: 10.1080/15248372.2012.742902

Butler, L. P., & Markman, E. M. (2014). Preschoolers use pedagogical cues to guide radical reorganization of category knowledge. *Cognition*, *130*(1), 116–127. doi: 10.1016/j.cognition.2013.10.002

Chalik, L., & Rhodes, M. (2018). Learning about social category-based obligations. *Cognitive Development*, *48*, 117–124. doi: 10.1016/j.cogdev.2018.06.010

Chalik, L., & Rhodes, M. (2020). Groups as moral boundaries: A developmental perspective. In *Advances in Child Development and Behavior* (Vol. 58, pp. 63–93). Elsevier. doi: 10.1016/bs.acdb.2020.01.003

Foster-Hanson, E., & Rhodes, M. (2019). Normative Social Role Concepts in Early Childhood. *Cognitive Science*, *43*(8), e12782. doi: 10.1111/cogs.12782

Gavrilets, S., & Richerson, P. J. (2017). Collective action and the evolution of social norm internalization. *Proceedings of the National Academy of Sciences*, *114*(23), 6068–6073. doi: 10.1073/pnas.1703857114

Gelman, S. A., Ware, E. A., & Kleinberg, F. (2010). Effects of Generic Language on Category Content and Structure. *Cognitive psychology*, *61*(3), 273–301. doi: 10.1016/j.cogpsych.2010.06.001

Kemp, C., Perfors, A., & Tenenbaum, J. B. (2007). Learning overhypotheses with hierarchical Bayesian models. *Developmental Science*, *10*(3), 307–321. doi: 10.1111/j.1467-7687.2007.00585.x

Lew-Levy, S., Lavi, N., Reckin, R., Cristóbal-Azkarate, J., & Ellis-Davies, K. (2018). How Do Hunter-Gatherer Children Learn Social and Gender Norms? A Meta-Ethnographic Review. *Cross-Cultural Research*, *52*(2), 213–255. doi: 10.1177/1069397117723552

Love, B. C., Medin, D. L., & Gureckis, T. M. (2004). SUSTAIN: A Network Model of Category Learning. *Psychological Review*, *111*(2), 309–332. doi: 10.1037/0033-295X.111.2.309

Macrae, C. N., Milne, A. B., & Bodenhausen, G. V. (1994). Stereotypes as Energy-Saving Devices: A Peek Inside the Cognitive Toolbox. *Journal of Personality and Social Psychology*, *66*(1), 37–47.

Roberts, S. O., Gelman, S. A., & Ho, A. K. (2017). So It Is, So It Shall Be: Group Regularities License Children's Prescriptive Judgments. *Cognitive Science*, *41*(S3), 576–600. doi: 10.1111/cogs.12443

Roberts, S. O., Ho, A. K., & Gelman, S. A. (2017). Group presence, category labels, and generic statements influence children to treat descriptive group regularities as prescriptive. *Journal of Experimental Child Psychology*, *158*, 19–31. doi: 10.1016/j.jecp.2016.11.013

Roccas, S., & Brewer, M. B. (2002). Social Identity Complexity. *Personality and Social Psychology Review*, *6*(2), 88–106. doi: 10.1207/S15327957PSPR0602_01

Taverna, A. S., Padilla, M. I., Baiocchi, M. C., & Peralta, O. A. (2021). Collaborative pedagogy: 3-year-olds bring pedagogical cues into alignment with analogical reasoning to extract generic knowledge. *European Journal of Psychology of Education*, *36*(2), 423–438. doi: 10.1007/s10212-020-00475-4

Tenenbaum, J. B., Kemp, C., Griffiths, T. L., & Goodman, N. D. (2011). How to Grow a Mind: Statistics, Structure, and Abstraction. *Science*, *331*(6022), 1279–1285. doi: 10.1126/science.1192788

Van Bavel, J. J., & Cunningham, W. A. (2012). A Social Identity Approach to Person Memory: Group Membership, Collective Identification, and Social Role Shape Attention and Memory. *Personality and Social Psychology Bulletin*, *38*(12), 1566–1578. doi: 10.1177/0146167212455829

Van Bavel, J. J., Packer, D. J., & Cunningham, W. A. (2008). The neural substrates of in-group bias: A functional magnetic resonance imaging investigation. *Psychological Science*, *19*(11), 1131–1139. doi: 10.1111/j.1467-9280.2008.02214.x

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*(11), 3343–3354. doi: 10.1162/jocn_a_00016