

Social Projection and Cognitive Differentiation Co-Explain Self-Enhancement and In-Group Favoritism

Alex Koch^a, Johanna Woitzel^b, and Russell Roberts^a

^aUniversity of Chicago, Chicago, USA; ^bRuhr University Bochum, Bochum, Germany

The Inductive Reasoning Model (IRM; Krueger et al., this issue) makes two assumptions. First, people's self-concept is positive; they are aware that they have more positive attributes than negative attributes. Second, when people are uncertain about another person's or group's attributes, they project their own attributes onto them because they are aware that social entities are similar to one another (i.e., share attributes). This social projection decreases with increasing social distance between the self and the other person or group. For example, people project their attributes onto in-groups members more than out-group members. These two assumptions have received ample empirical support (Robbins & Krueger, 2005; Unkelbach et al., 2019). The IRM combines self-positivity and differential social projection to predict the size of four effects in the social psychological literature: self-enhancement, in-group favoritism, intergroup accentuation, and differential accuracy (i.e., a more accurate description of in-groups vs. out-groups). Further, the IRM predicts the correlations between these effect sizes from self-positivity and differential social projection. In sum, the model pinpoints the (co)variation of important effects by relying on the evaluative information ecology and a social-cognitive heuristic instead of the motivational processes emphasized by theories of social identity (Ellemers et al., 2002) and social comparison (Mussweiler, 2003). We appreciate the IRM's parsimony and generativity, and we draw on our and other previous research to discuss its validity and generality by focusing on a related cognitive-ecological model of impression formation.

First, we encourage studies that test the validity of the IRM. To establish internal validity, studies may examine whether varying the modeled causes or inputs moderates the four effects in ways that are consistent with the model (Krueger et al., this issue). A neutral self-concept or indiscriminate social projection between in-group and out-group targets should result in null effects that are less informative than systematic differences (but see Lakens et al., 2020). Thus, we suggest reversing or amplifying the effects. A first study may manipulate the valence of people's self-concept (e.g., through upward social comparison or false feedback on a bogus test). When people's self-concept is negative, the IRM predicts self-effacement, out-group favoritism, and differential inaccuracy (i.e., a less accurate description of in-groups vs. out-groups). In a second study, people whose

self-concept is positive may encounter an out-group that is "similar to the self" versus an in-group that "differs from the self." In this unusual ecology, self-aggrandizement (i.e., exaggerated self-enhancement), out-group favoritism, and differential inaccuracy should emerge, according to the IRM. As implied by Krueger et al. (this issue), a third study may cross manipulations of self-concept valence and differential social projection to test whether these variables interact to predict the IRM's outputs.

To establish external validity, another study could predict holistic impressions from the IRM's computations. For example, the IRM computes a person's self-enhancement by subtracting a first correlation over attributes from a second correlation over attributes. The first correlation is between a person's ratings of their in-group having the attributes and the attributes' desirability. The second correlation is between the person's ratings of the self having the attributes and, again, the attributes' desirability. However, people's mental capacity is limited and they need to represent information efficiently. Hence, they form influential overall impressions of (the positivity of) the self and others (Abele et al., 2021; Koch et al., 2021). Thus, it seems important that the person's self-enhancement as computed in the IRM predicts the difference between their overall impressions of the self and their in-group. In a similar vein, future research could test whether in-group favoritism and intergroup accentuation, as computed in the IRM, predict the difference between people's overall impression of the in-group and out-group, their holistic impression of group-level diversity, and the difference between the perceived accuracy of their characterizations of the in-group and out-group.

Second, we encourage studies that examine the generality of the IRM. One direction we consider worthy of exploration relies on the four types of psychological distance postulated by Construal Level Theory (CLT; Trope & Liberman, 2010). CLT argues that the social distance from the self increases as people reason about an out-group, compared to an in-group. Distance from the self also increases, however, as the same social entity is further away in space, time, or certainty of interaction. All four types of distance may decrease social projection, and thereby moderate the IRM's effects in predicted ways. For example, the IRM may predict that people favor not just in-groups over out-groups,

but also one and the same group that is closer to the self in space, time, or probability of social interaction.

Social entities can differ from one another in many ways, some of which may be more relevant to differential social projection. Prior research prompted people to sort society-representative samples of groups according to (dis)similarity without prompts on how to interpret the concept of (dis)similarity. People spontaneously differentiated the groups according to their socioeconomic status and ideological beliefs (Koch et al., 2020; Koch et al., 2022; Nicolas et al., 2022, see also Imhoff et al., 2018; Koch et al., 2018). Thus, social projection may be more differential when the difference between a person's in-group and out-group is a status gap or an ideological divide, compared to another type of social distance. As a result, the IRM may predict larger effect sizes for self-enhancement, in-group favoritism, intergroup accentuation, and differential accuracy when the difference between the person's in-group and out-group is a status gap or an ideological divide. This would be consistent with a growing body of evidence that status and ideology predict where people live and whom they communicate, cooperate, and mate with (Gentzkow & Shapiro, 2011; Roberts & Koch, 2024; Woitzel & Koch, 2023).

A challenge to the generality of the IRM is that people often know the attributes of others because they have encountered them, leaving little room for social projection and requiring other explanations of self-enhancement, in-group favoritism, etc. The cognitive-ecological model (CEM; Alves et al., 2017a, 2023) is similar to the IRM in that it predicts self-enhancement, etc., without relying on motivational processes. According to the CEM, impression formation is often a serial process in which social targets are encountered individually over time. Examples of this include job interviews throughout a business day or browsing through profiles on a dating app. In serial impression formation, perceivers tend to overlook, ignore, and forget a target's attributes if they were already present in previous targets (Kruschke, 2003; Woitzel & Alves, 2024). That is, people process the unique, differentiating attributes of new encounters in a series of targets (Bassok & Trope, 1983–1984; Hamilton & Gifford, 1976; Reeder & Brewer, 1979).

The claims of the CEM are rooted in two empirically supported asymmetries of information distribution: diversity asymmetry and frequency asymmetry. Diversity asymmetry is the idea that there are more negative (vs. positive) attributes that a target *can* have; frequency asymmetry is the idea that the typical target *has* more positive (vs. negative) attributes. Thus, the CEM agrees with the IRM's assumption that most people's self-concept is positive. The diversity and frequency asymmetries result in two relevant aspects of the evaluative information ecology (EvIE; Unkelbach et al., 2019). First, the attributes that targets *share* are more likely to be positive (Alves et al., 2016). Second, the distinct attributes that make targets *unique* are more likely to be negative (Alves et al., 2017a, 2017b). Thus, when people form impressions of others, they focus on attributes that allow them to distinguish focal targets from those they have encountered previously, and these attributes tend to

be more likely to be negative for targets who appear later in the series. Thus, their overall impressions of later-encountered targets are more likely to be negative (e.g., Wang et al., 2023).

A recent paper confirmed this indirect effect from serial position to differentiation to negativity in natural language (i.e., spontaneous) descriptions. Notably, the randomly ordered targets were real individuals as they appeared in their profile picture on the world's largest social media platform or in a popular reality TV show (Koch et al., 2024). Consistent with the CEM, the size of the effect was largest when the targets had many positive attributes that they shared with one another. In contrast, a serial position-positivity effect emerged in a non-representative ecology where the targets had many negative attributes that they shared with one another. In two non-representative ecologies where the targets' positive or negative attributes were distinct/unique, serial position had no effect on the algorithmic and subjective valence of the perceiver's descriptions of the targets. Other research confirmed the evaluative disadvantage of appearing later in a series in hindsight descriptions and overall impressions of groups that had shared positive and distinct negative attributes (Alves et al., 2018; Woitzel & Alves, 2024).

The serial position-negativity effect is relevant for the IRM because it is reasonable to assume that social distance correlates with serial position in social interaction as well as impression formation. People are with themselves from the start of their lives, and they likely get to know and interact with their in-groups before out-groups (Sherman et al., 2009). Thus, more often than not, people may first assess the attributes of the self, then their in-groups, and only then out-groups. In the real-world ecology where people's shared and unique attributes are positive and negative, respectively, the aforementioned order of targets predicts that cognitive differentiation will reduce the likelihood that the perceiver acknowledges the presence of positive attributes in out-groups (vs. in-groups), and in in-groups (vs. the self). Conversely, cognitive differentiation will increase the likelihood that the perceiver acknowledges the presence of negative attributes in out-groups (vs. in-groups), and in in-groups (vs. the self). As a result, the CEM can explain self-enhancement, in-group favoritism, intergroup accentuation, and differential accuracy in contexts where the perceiver's uncertainty about these targets' attributes is rather low because impressions are formed (serially) based on (serial, fixed-order) interactions with the in-group and out-group targets.

An interesting avenue for future research is to disentangle whether, in a context with high uncertainty about the attributes of minimal groups, self-enhancement, in-group favoritism, intergroup accentuation, and differential accuracy emerge due to social projection decreasing with social distance, as argued by the IRM, or serial differentiation (independent of social distance) as argued by the CEM (Alves et al., 2023). The meta-analysis by Robbins and Krueger (2005) found a small but non-significant decrease in social projection when study participants rated the attributes of

the self before rating the attributes of groups (vs. the other way round). Still, varying the serial order of rating the attributes of the self, in-group, and out-group in high-powered experiments seems worthy of examination. Another way to pit the models against each other would be to have participants rate the attributes of the self, then several in-groups, and finally several out-groups. The IRM predicts less favorable attitudes toward the out-groups (vs. in-groups), but no change in the attitudes toward different in-groups, and different out-groups. If serial position matters as predicted by the CEM, however, attitudes should be more favorable toward both those in-groups and out-groups that people rate earlier, compared to later.

Disclosure Statement

No potential conflict of interest was reported by the author(s).

References

- Abele, A., Ellemers, N., Fiske, S., Koch, A., & Yzerbyt, V. (2021). Navigating the social world: Shared horizontal and vertical evaluative dimensions. *Psychological Review*, *128*(2), 290–314. <https://doi.org/10.1037/rev0000262>
- Alves, H., Koch, A., & Unkelbach, C. (2016). My friends are all alike—the relation between liking and perceived similarity in person perception. *Journal of Experimental Social Psychology*, *62*, 103–117. <https://doi.org/10.1016/j.jesp.2015.10.011>
- Alves, H., Koch, A., & Unkelbach, C. (2017a). Why good is more alike than bad: Processing implications. *Trends in Cognitive Sciences*, *21*(2), 69–79. <https://doi.org/10.1016/j.tics.2016.12.006>
- Alves, H., Koch, A., & Unkelbach, C. (2017b). The “common good” phenomenon: Why similarities are positive and differences are negative. *Journal of Experimental Psychology: General*, *146*(4), 512–528. <https://doi.org/10.1037/xge0000276>
- Alves, H., Koch, A., & Unkelbach, C. (2018). A cognitive-ecological explanation of intergroup biases. *Psychological Science*, *29*(7), 1126–1133. <https://doi.org/10.1177/0956797618756862>
- Alves, H., Koch, A., & Unkelbach, C. (2023). Evaluative consequences of sampling distinct information. In K. Fiedler, P. Juslin, & J. Denrell (Eds.), *Sampling in Judgment and Decision Making* (pp. 222–244). Cambridge University Press.
- Bassok, M., & Trope, Y. (1983–1984). People’s strategies for testing hypotheses about another’s personality: Confirmatory or diagnostic? *Social Cognition*, *2*(3), 199–216. <https://doi.org/10.1521/soco.1984.2.3.199>
- Ellemers, N., Spears, R., & Doosje, B. (2002). Self and social identity. *Annual Review of Psychology*, *53*(1), 161–186. <https://doi.org/10.1146/annurev.psych.53.100901.135228>
- Gentzkow, M., & Shapiro, J. M. (2011). Ideological segregation online and offline. *The Quarterly Journal of Economics*, *126*(4), 1799–1839. <https://doi.org/10.1093/qje/qjr044>
- Hamilton, D. L., & Gifford, R. K. (1976). Illusory correlation in interpersonal perception: A cognitive basis of stereotypic judgments. *Journal of Experimental Social Psychology*, *12*(4), 392–407. [https://doi.org/10.1016/S0022-1031\(76\)80006-6](https://doi.org/10.1016/S0022-1031(76)80006-6)
- Imhoff, R., Koch, A., & Flade, F. (2018). (Pre)occupations: A data-driven map of jobs and its consequences for categorization and evaluation. *Journal of Experimental Social Psychology*, *77*, 76–88. <https://doi.org/10.1016/j.jesp.2018.04.001>
- Koch, A., Imhoff, R., Unkelbach, C., Nicolas, G., Fiske, S., Terache, J., Carrier, A., & Yzerbyt, V. (2020). Groups’ warmth is a personal matter: Understanding consensus on stereotype dimensions reconciles adversarial models of social evaluation. *Journal of Experimental Social Psychology*, *89*, 103995. <https://doi.org/10.1016/j.jesp.2020.103995>
- Koch, A., Bromley, A., Woitzel, J., & Alves, H. (2024). Differentiation in social perception: Why later-encountered individuals are described more negatively. *Journal of Personality and Social Psychology*. Advance online publication. <https://doi.org/10.1037/pspa0000383>
- Koch, A., Kervyn, N., Kervyn, M., & Imhoff, R. (2018). Studying the cognitive map of the U.S. States: Ideology and prosperity stereotypes predict interstate prejudice. *Social Psychological and Personality Science*, *9*(5), 530–538. <https://doi.org/10.1177/1948550617715070>
- Koch, A., Speckmann, F., & Unkelbach, C. (2022). Q-SpAM: Using spatial arrangement to measure similarity in efficient online research powered by Qualtrics. *Sociological Methods & Research*, *51*(3), 1442–1464. <https://doi.org/10.1177/0049124120914937>
- Koch, A., Yzerbyt, V., Abele, A., Ellemers, N., & Fiske, S. T. (2021). Social evaluation: Comparing models across interpersonal, intra-group, intergroup, several-group, and many-group contexts. In B. Gawronski (Ed.), *Advances in Experimental Social Psychology* (pp. 1–68). Academic Press.
- Kruschke, J. K. (2003). Attention in learning. *Current Directions in Psychological Science*, *12*(5), 171–175. <https://doi.org/10.1111/1467-8721.01254>
- Lakens, D., McLatchie, N., Isager, P. M., Scheel, A. M., & Dienes, Z. (2020). Improving inferences about null effects with Bayes factors and equivalence tests. *The Journals of Gerontology: Series B*, *75*(1), 45–57. <https://doi.org/10.1093/geronb/gby065>
- Mussweiler, T. (2003). Comparison processes in social judgment: Mechanisms and consequences. *Psychological Review*, *110*(3), 472–489. <https://doi.org/10.1037/0033-295X.110.3.472>
- Nicolas, G., Fiske, S. T., Koch, A., Imhoff, R., Unkelbach, C., Terache, J., Carrier, A., & Yzerbyt, V. (2022). Relational versus structural goals prioritize different social information. *Journal of Personality and Social Psychology*, *122*(4), 659–682. <https://doi.org/10.1037/pspi0000366>
- Reeder, G. D., & Brewer, M. B. (1979). A schematic model of dispositional attribution in interpersonal perception. *Psychological Review*, *86*(1), 61–79. <https://doi.org/10.1037/0033-295X.86.1.61>
- Roberts, R., & Koch, A. (2024). Power polarizes moral evaluations. *Personality & Social Psychology Bulletin*. Advance online publication. <https://doi.org/10.1177/01461672241245181>
- Robbins, J. M., & Krueger, J. I. (2005). Social Projection to Ingroups and Outgroups: A Review and Meta-Analysis. *Personality and Social Psychology Review: An Official Journal of the Society for Personality and Social Psychology, Inc.*, *9*(1), 32–47. https://doi.org/10.1207/s15327957pspr0901_3
- Sherman, J. W., Kruschke, J. K., Sherman, S. J., Percy, E. J., Petrocelli, J. V., & Conrey, F. R. (2009). Attentional processes in stereotype formation: A common model for category accentuation and illusory correlation. *Journal of Personality and Social Psychology*, *96*(2), 305–323. <https://doi.org/10.1037/a0013778>
- Trope, Y., & Liberman, N. (2010). Construal-level theory of psychological distance. *Psychological Review*, *117*(2), 440–463. <https://doi.org/10.1037/a0018963>
- Unkelbach, C., Koch, A., & Alves, H. (2019). The evaluative information ecology: On the frequency and diversity of “good” and “bad”. *European Review of Social Psychology*, *30*(1), 216–270. <https://doi.org/10.1080/10463283.2019.1688474>
- Wang, J., Chaudhry, S., & Koch, A. (2023). Reminders undermine impressions of genuine gratitude. *Journal of Personality and Social Psychology*. Advance online publication. <https://doi.org/10.1037/pspi0000442>
- Woitzel, J., & Alves, H. (2024). The formation of negative attitudes towards novel groups. *Psychological Science*. Advance online publication. <https://doi.org/10.1177/09567976241239932>
- Woitzel, J., & Koch, A. (2023). Ideological prejudice is stronger in ideological extremists (vs. moderates). *Group Processes & Interpersonal Relations*, *26*(8), 1685–1705. <https://doi.org/10.1177/13684302221135083>