The First-Member Heuristic: Group Members Labeled "First" Influence Judgment and Treatment of Groups

Date submitted: May 6th 2019

People often make judgments about a group (e.g., immigrants from a specific country) based on information about a single group member. Seven studies (N = 1,929) tested the hypothesis that people will expect the performance of an arbitrarily ordered group to match that of the group member in the first position of a sequence more closely than that of group members in other positions. This greater perceived diagnosticity of the first member will in turn affect how people treat the group. This pattern of judgment and treatment of groups, labeled the "first-member heuristic," generalized across various performance contexts (e.g., gymnastic routine, relay race, job performance), and regardless of whether the focal member performed poorly or well (Studies 1-3). Consistent with the notion that first members are deemed most informative, participants were most likely to turn to the member in the first (vs. other) position to learn about the group (Study 4). Further, through their disproportionate influence on the expected performance of other group members, first members' performances also influenced participants' support of policies that would benefit or hurt a group (Study 5) and their likelihood to join a group (Study 6). Finally, perceived group homogeneity moderated the first-member heuristic, such that it attenuated for nonhomogeneous groups (Study 7).

Keywords: first-member heuristic; diagnosticity; group perception; sequences; order effects

People often make judgments about groups (e.g., immigrants from a specific country) based on information about a single focal member of the group. In these cases, representations of a specific individual (an exemplar) influence judgments about others deemed to be from the same group (Smith & Zarate, 1992). In particular, group membership often determines the extent to which people see one person as diagnostic of other people—that is, as informative for drawing inferences about others. For example, after encountering a visitor from a distant country who happens to be a skilled tennis-player, one might infer that residents of that country (i.e., the visitor's group) tend to be good at tennis. In this case, the target person appears diagnostic of a group (Hamilton et al., 2015).

What factors might determine the extent to which people use a particular group member to draw inferences about other group members? We propose that a group member's position in an ordered sequence of group members might matter. A characteristic of everyday groups is that they are often organized and numbered sequentially (Bar-Hillel, 2015). These sequential orders can carry meaning, such that the first employee of a company might have—directly or indirectly—influenced subsequent hiring decisions, and hence might provide a good indication of the profile of future employees. Yet the sequential orders assigned to group members can also be completely arbitrary. For example, an arbitrary factor, such as the first letter of students' last names, can determine the order in which they present their final class projects. In this case, the order (e.g., first versus tenth student to present) carries no meaning, because a student whose last name starts with an "A" should provide no more information about the group of students presenting their projects than a student whose name starts with an "S." Similarly, reviewers of manuscripts for academic journals, or cash registers in stores are often ordered arbitrarily (e.g., reviewer 1, cashier number 2).

In the present research, we explore whether the position of a focal group member in an arbitrarily ordered group influences the extent to which people perceive this group member as diagnostic of the group. We define a group as a set of items that are put together in a mental category based on some shared characteristic(s). We argue that people overgeneralize from situations in which the sequential order of group members carries meaning (e.g., the first employee of a company having shaped the subsequent hiring of employees) to situations in which this order is purely arbitrary. Specifically, we propose that people will perceive a member in the first position to be more diagnostic of the group than members in other positions (e.g., middle or last), such that they will weigh the performance of this first (vs. any subsequent) member more heavily in their expectations and predictions about the rest of the group. In turn, we expect these judgments to influence how people treat and behave toward the group.

Diagnosticity in Judgment

The diagnosticity of a piece of information or cue (e.g., cooks for the homeless every day after work) about an attitude object (e.g., Sue) refers to the extent to which this cue increases the probability that the object belongs to one category (e.g., Sue is generous) and lowers the probability that the object belongs to any other category (e.g., Sue is selfish; Skowronski & Carlston, 1987). Hence, diagnostic cues tend to receive more weight (i.e., are more influential) in judgment than non-diagnostic cues; they are deemed more informative (Birnbaum, 1972; Cone & Ferguson, 2015; Nisbett, Zukier, & Lemley, 1981; Menon, Raghubir, & Schwarz, 1995; Risen, Gilovich, & Dunning, 2007; Skowronski, 2002; Wyer & Hinkle, 1976). For example, people are more likely to draw inferences about others based on extreme behaviors than based on moderate ones (Reeder, 1985; Reeder & Brewer, 1979). Indeed, an extreme behavior such as jumping off a high cliff into a lake would seem indicative of an extreme thrill-seeker, whereas a

relatively moderate behavior such as jumping off a low rock into a lake could indicate both extreme thrill-seeking and fun-loving personalities.

Just as specific characteristics can seem diagnostic of a person, a person can seem diagnostic or informative of other people (Hamilton et al., 2015). According to Smith and Zarate's (1992) exemplar-based model of social judgment, when an exemplar seems particularly diagnostic, people will use their knowledge about her to make inferences about other group members. In particular, a salient, often irrelevant, similarity between two individuals (e.g., race, gender, age, or physical appearance) can determine the extent to which people see one individual as diagnostic of the other (White & Shapiro, 1987). For example, in one study, Gilovich (1981) gave participants—all college football coaches and sportswriters—brief descriptions of fictitious college football players. Some descriptions created an irrelevant similarity between the described college player and a successful professional player (e.g., the college player and the pro came from the same hometown), whereas control descriptions offered no such similarity. Participants rated the college players with similarity to a famous professional player (vs. control) as having significantly higher potential for future success, which indicated that they judged the professional player as diagnostic for making inferences about the college player.

In a group context, the perceived similarity among group members also influences the extent to which each group member is deemed diagnostic of the rest of the group (Campbell, 1958; Lickel et al., 2000). Crawford, Sherman, and Hamilton (2002) found that when people encounter an individual who belongs to a group that is perceived to form a single and coherent entity (i.e., an entitative group), they transfer their impression of that group member to the entire group. Thus, if one member of such a group is described as intelligent, this trait is attributed to the entire group. Traits and characteristics of the group as a whole are subsequently transmitted

to the group's individual members, such that an unknown member of the particular group would also be expected to be intelligent (see also Hamilton, Sherman, & Castelli, 2002; Spencer-Rodgers, Williams, Hamilton, Peng, & Wang, 2007; Susskind, Maurer, Thakkar, Hamilton, & Sherman, 1999). Thus, the diagnosticity of group member 'A' for drawing inferences about group member 'B' can stem from the extent to which these two group members seem to share common characteristics (i.e., are similar to each other).

Beyond similarity, what other factors might determine a group member's diagnosticity for making inferences about the rest of the group? We explore whether the position of a group member (e.g., first, middle, last) in an arbitrarily ordered group affects diagnosticity judgments. Indeed, groups are often organized sequentially (e.g., students on a class list, cash registers in a store), and this sequence might determine the order in which group members perform a specific task or activity (e.g., gymnastic performances at the Olympics, singing contests, or students defending their theses). We investigate the possibility that the performance of a focal group member will be deemed more diagnostic of the rest of the group if this focal member is in the first (vs. any subsequent) position in the group sequence.

The Peculiarities of the First

Previous research documented that people perceive the first position in a sequence differently from other positions (Anderson, 1973; Mantonakis, Rodero, Lesschaeve, & Hastie, 2009; Shteingart, Neiman, & Loewenstein, 2013). For example, the primacy effect in memory refers to the tendency to remember the first stimulus in a sequence better than the ones in the middle (Glanzer & Cunitz, 1966; Murdock, 1962; Sulmont-Rosse, Chabanet, Issanchou, & Köster, 2008). First information also dominates impression formation by its larger impact on the overall impression people form of others, relative to subsequent information (Anderson, 1973;

Sullivan, 2019). For example, people perceive someone described as “friendly” and “stubborn” more positively than someone described as “stubborn” and “friendly.” Indeed, when people process the positive attribute (“friendly”) first, they then process subsequent negative attributes (“stubborn”) in light of the initial positive information. Furthermore, people typically pay more attention to the first piece of information than to information that comes later (Holyoak & Simon, 1999; Sulmont-Rosse et al., 2008).

People not only remember the first position better, but they also tend to show a preference for the first person or object in a sequence. Presumably, people compare each subsequent person or object to the first, which they remember in more detail. Furthermore, the first tends to leave the strongest impression (MacFie, Bratchell, Greenhoff, & Vallis, 1989; Page & Page, 2010; Pandelaere, Millet, & Van den Bergh, 2010). Aside from leaving a stronger impression and memory trace, objects in first positions are often deemed to be the best (Carney & Banaji, 2012), and in some cases people infer they are the cause of what comes next (Einhorn & Hogarth, 1986).

Underlining that people might infer influence from first positions, LeBoeuf, Williams, and Brenner (2014) found that merely labeling an experience as first increases its influence on expectations of future experiences—even when this ordering is largely irrelevant for the judgment at hand. These arbitrarily labeled “phantom firsts” influenced, for example, expectations about the weather. Participants who learned about the weather in Paris—five sunny days followed by one rainy day—expected the weather in Paris to be rainier if the single rainy day happened to be their first (vs. any other) vacation day in Paris. Phantom firsts are perceived to foreshadow future events because they create temporal markers, such that what happened before the marker becomes less relevant, and the marker (i.e., the first event) becomes

particularly diagnostic of what is to come next (see also, Dai, Milkman, & Riis, 2014; Peetz & Wilson, 2013; Zhao, Lee, & Soman, 2012).

In sum, the literature on judgments about sequential positions shows that first positions are preferred and remembered better relative to other positions. Notably, prior research examined judgments about objects that are presented sequentially (Page & Page, 2010; Sulmont-Rosse et al., 2008) or predictions about future experiences based on information that is sequentially ordered (LeBoeuf et al., 2014). In this previous work, people learned information about several people, objects, or experiences from a particular sequence. In contrast, we investigate contexts in which people receive information about the performance of a single member of an arbitrarily ordered group (e.g., the one in the first, middle, or last position). We test the extent to which people rely on this information to draw inferences about the other members of the group. In particular, we explore whether first positions are perceived differently in settings where they cannot foreshadow other positions or cannot be compared to later positions.

The First-Member Heuristic

We study the "first-member heuristic." Thereby, we refer to the tendency to make judgments, and in turn treat group members, based on information on a first (vs. other) group member in an arbitrarily ordered sequence.

At times, the first member of a group is objectively more informative or diagnostic of the group than subsequent members. Indeed, in many contexts, the first member of a group either created the group or served as the foundation for creating the group, such that the other members of the group were selected because of relevant shared features with the first. For example, if Sue wants to start a running club, she will most likely look for runners at a similar fitness level as herself, and possibly of the same gender and age group. If Al wants to start a basketball team

including his 10-year-old daughter, he will most likely seek out parents of other children in this age group and residing within the same area. Similarly, a meeting with the first employee of a new start-up company might be particularly telling of the company's corporate culture in terms of the types of qualifications and personality types the company's founders intend to hire in the future. Additionally, this first employee might have some input into the hiring of subsequent employees, which most likely would result in hiring people similar to him or her (McPherson, Smith-Lovin, & Cook, 2001). Finally, the first member might exert more influence on subsequent members, for example by setting behavioral norms that subsequent members will follow. Thus, for various reasons, the first member of a group might provide a good indication of the profile of subsequent members.

We propose that this learned informative value of first members—persons, animals, or objects in first positions—will extend to contexts in which it would not be logical to expect a group member to be particularly informative merely because this member is in the first position. Specifically, we expect people to apply the learned informative value of first members—i.e., perceiving a group member in the first (vs. any subsequent) position as more diagnostic of the rest of the group—even when the first member is randomly determined to be in that position and cannot influence subsequent members. Such overgeneralizations can be found in other domains of social judgment and behavior (e.g., Rand et al., 2014; Yamagishi, Terai, Kiyonari, Mifune, & Kanazawa, 2007). For example, cooperation is often an advantageous strategy. Thus, people form an intuition to cooperate, which is then overgeneralized to situations in which cooperation is not advantageous (e.g., one-shot anonymous economic games).

Our first hypothesis is that there will be greater consistency between the performance of the first (vs. any subsequent) member and the expected performance of the rest of the group.

Thus, when the first (vs. any subsequent) member performs well, people will expect the other group members to perform better, and when the first member (vs. any subsequent) member performs poorly, people will expect the other group members to perform worse. Furthermore, we predict that when trying to evaluate or predict a group's performance, people will be more likely to examine the performance of the first (vs. any subsequent) member of the group, expecting this member's performance to be the most informative of the whole.

The notion that people will rely more heavily on the first member of an (arbitrarily) ordered group to make judgments about the rest of the group has implications for their behavior toward the group. Specifically, our second hypothesis is that given its disproportionate weight on group evaluations, the performance of the group member in the first (vs. any subsequent) position will influence people's treatment of the group in relevant performance contexts such as (A) supporting policies that benefit or hurt the group, and (B) joining the group. For example, upon hearing that the first (vs. any subsequent) member of an arbitrarily ordered group of immigrants entering the United States performed poorly at his job, people will expect the rest of the group to perform worse, and hence will be more likely to support policies against this group of immigrants. Thus, we propose that people's judgment and treatment of (arbitrarily ordered) groups will be more dependent upon their beliefs about the first (vs. any subsequent) member of the group. We label this pattern of judgment and behavior the "first-member heuristic."

Finally, we expect the first-member heuristic to occur in the evaluation of groups—rather than the evaluation of any collection of ordered items or broad category of items. Indeed, perceptions of similarity along salient dimensions (i.e., homogeneity) play an important role not only in the perception that a set of items (people, animals, objects) forms a group, but also in the perception that each item can be diagnostic of the rest (Hamilton et al., 2015; Smith & Zarate,

1992). A collection of items that is low in homogeneity should be less likely to be considered a group, and the perceived diagnosticity of each member should be low. Thus, our third hypothesis is that group homogeneity will moderate the first-member heuristic, such that it will attenuate for groups perceived as low in homogeneity (i.e., heterogeneous group).

We report seven studies that tested our hypotheses about the first-member heuristic. Across these studies, we estimated a minimum required sample size of 40 participants per experimental condition to obtain a desired power of .80 at a desired alpha level of $p = .05$, and average effect sizes of $r = .32$ (*Cohen's d* = .66) similar to those documented in previous research on order effects (Carney & Banaji, 2012). Several studies exceeded the minimum sample sizes of 40 participants per condition, which we explain in the respective method sections. We report all measures and manipulations in all studies. Survey materials and datasets are available at https://tinyurl.com/first-member.

Study 1: Predicting the Performance of Group Members

Study 1 tested the first-member heuristic in judgments of an arbitrarily ordered group of gymnasts. Participants learned about the good performance of one gymnast and indicated their expectations about the performance of the other group members. To eliminate any inferences based on conversational norms, we emphasized that the computer would select the group member about whose performance participants would learn.[1] We predicted that learning about

[1] Conversational norms—as theorized by Grice (1957, 1975)—typically dictate that information be discussed in sequential or chronological order (the Gricean "maxim of manner"). Thus, when participants are told about the performance of the first member, they might assume the information is just being presented sequentially. Within this perspective, there should be nothing special or different about the first, and hence participants should have no reason to expect others to have performed any better or worse than the first. However, when participants are told only about a group member in another sequential position (e.g., middle or last), they might assume there is something noteworthy or special about this member's performance that justifies skipping the preceding members—that is, breaking the maxim of manner. Hence, they would expect other

the good performance of the first (vs. middle or last) group member would raise participants' expectations for the rest of the group.

Method

Participants and Design. We opened the study to 300 participants on Amazon's Mechanical Turk (MTurk), and 305 participants (162 females; M_{age} = 37.70, SD = 11.97) completed the study on desktop or laptop computers (no mobile phones) in exchange for $0.30. In this study, we recruited 100 participants per condition to provide a high-powered test of our proposed effect. The study employed a one-factorial (position of focal group member: first vs. middle vs. last) between-subjects design. Another six participants quit the survey after condition assignment but before completing the dependent variables (one from the first condition, two from the middle condition, and three from the last condition).

Materials and Procedure. Participants read about a group of seven college gymnasts who were performing a floor-routine sequentially (presumably, in order to be considered for a national competition). The scenario emphasized that the performance order was determined by a lottery (each gymnast draw a number out of a jar), to eliminate inferences about the gymnast's motivation or talent based on her sequential position.

Participants further read that the computer would select one group member and would share her performance with the participant. To emphasize this arbitrary computerized selection, participants had to wait for a few seconds while the computer supposedly selected a gymnast. Critically, participants then read that Emma, the first, fourth (i.e., middle), or last gymnast to

---

group members to have performed differently from this member (better or worse). Indeed, conversational norms also dictate that discussions include the most relevant and pertinent information (the Gricean "maxim of relation"). In the present study, emphasizing the arbitrary computer selection allowed us to preclude such inferences due to conversational norms.

perform (depending on the condition) gave a stunning performance and that her routine was flawless.

To assess expectations about the group's performance, we asked participants to indicate how they expected each of the other gymnasts —Jenny, Katie, Sofia, Ava, Gaby, and Lisa— had performed (six ratings; 1 = *very poorly*; 9 = *very well*). Because participants received no individuating information about the other gymnasts, we expected their assessment to be similar across gymnasts. Thus, we collapsed participants' rating into an index of expected group-performance ($\alpha = .89$).

Finally, as an attention check, we asked participants how the order in which the gymnasts performed had been determined (1 = *The best gymnast performed first*; 2 = *The gymnasts picked numbers out of a jar*; 3 = *The gymnast chose the order among themselves*; 4 = *I don't remember*). Thirteen percent of participants falsely reported that the best gymnast had gone first (option "1") and thus failed the attention check (18 from the first condition, nine from the middle condition, and 13 from the last condition). Another eight percent of participants reported that they did not remember (option "4") how the order had been assigned (eight from the first condition, six from the middle condition, and nine from the last condition). We retained all participants who failed the attention check and those who did not remember the order assignment in the subsequent analyses and note that excluding them did not change the pattern or the significance of the results.

Results and Discussion

In support of the hypothesis about the first-member heuristic, expected group-performance varied based on the focal member's position, $F(2, 303) = 5.957$, $p = .003$, $\eta^2 = .04$. Specifically, when the high performer was in the first position, participants expected the rest of

the group to have performed better (M = 6.98, SD = 1.29) than when the high performer was in the middle (M = 6.46, SD = 1.27), p = .003, CI[0.173, 0.869][2] or last position (M = 6.45, SD = 1.22), p = .003, CI[0.188, 0.884]. There was no difference in expected group-performance between the middle and last conditions, p = .934, CI[-0.333, 0.363].

This study provided initial evidence for the first-member heuristic: positive information about the first (vs. middle or last) group member was more influential for judgments of other group members. Because in this study we only compared the first to the middle and last member, we went on to compare the first member to any other position in the group in the next study.

Study 2: Comparing the First Member to All Subsequent Members

Study 2 tested the first-member heuristic by comparing performance expectations following information about the first member of a group versus every other group member. We predicted that the first group member would be deemed more diagnostic for inferences about the group than any subsequent member.

Method

Participants and Design. We opened the study to 500 participants on MTurk, and 504 participants (254 females; M_{age} = 35.42, SD = 11.51) completed the study on desktop or laptop computers (no mobile phones) in exchange for $0.35. The sample size (100 participants per condition) was similar to that of Study 1. The study employed a one-factorial (position of focal group member: first vs. second vs. third vs. fourth vs. fifth) between-subjects design. Another 10 participants quit the survey after condition assignment but before completing the dependent variables (two from the first condition, one from the second condition, four from the third

[2] We report 95% confidence intervals for the mean difference unless otherwise specified.

condition, two from the fourth condition, and one from the fifth condition). We pre-registered the data collection and analysis plan for this study at http://aspredicted.org/blind.php?x=yf2td7.

Materials and Procedure. Participants read the same scenario about a group of gymnasts as in Study 1, except that the group in this study consisted of five members. The focal gymnasts, Emma, was assigned by random draw to go first, second, third, fourth, or fifth (depending on the condition). Participants further completed the same dependent variable and responded to the same attention check. Eleven percent of participants falsely reported that the best gymnast had gone first and thus failed the attention check (10 from the first, second, third, and fifth conditions, and 13 from the fourth condition). Another 10 percent of participants reported that they did not remember how the order had been assigned (13 from the first condition, 10 from the second condition, six from the third condition, 12 from the fourth condition, and 10 from the fifth condition). We retained both the participants who failed the attention check and those who did not remember the order assignment in the subsequent analyses and note that excluding them did not change the pattern or the significance of the results.

Results and Discussion

In support of the hypothesis about the first-member heuristic, expected group-performance varied based on the focal member's position, $F(4, 500) = 10.971$, $p < .001$, $\eta^2 = .08$ (see Figure 1). Specifically, when the high performer was in the first position, participants expected the rest of the group to have performed better ($M = 7.07$, $SD = 1.32$) than when she was second ($M = 6.39$, $SD = 1.43$), $p < .001$, CI[0.314, 1.045], third ($M = 5.98$, $SD = 1.21$), $p < .001$, CI[0.726, 1.463], fourth ($M = 6.03$, $SD = 1.36$), $p < .001$, CI[0.678, 1.411], or fifth ($M = 6.33$, $SD = 1.30$), $p < .001$, CI[0.378, 1.109]. The results also showed that when the high performer was second, the rest of the group was seen as better than when she was third, $p = .027$, CI[0.047,

0.782], or fourth, p = .050, CI[-0.001, 0.730]. Despite these two (unpredicted) differences, the first condition was the only one that differed significantly from all other conditions.

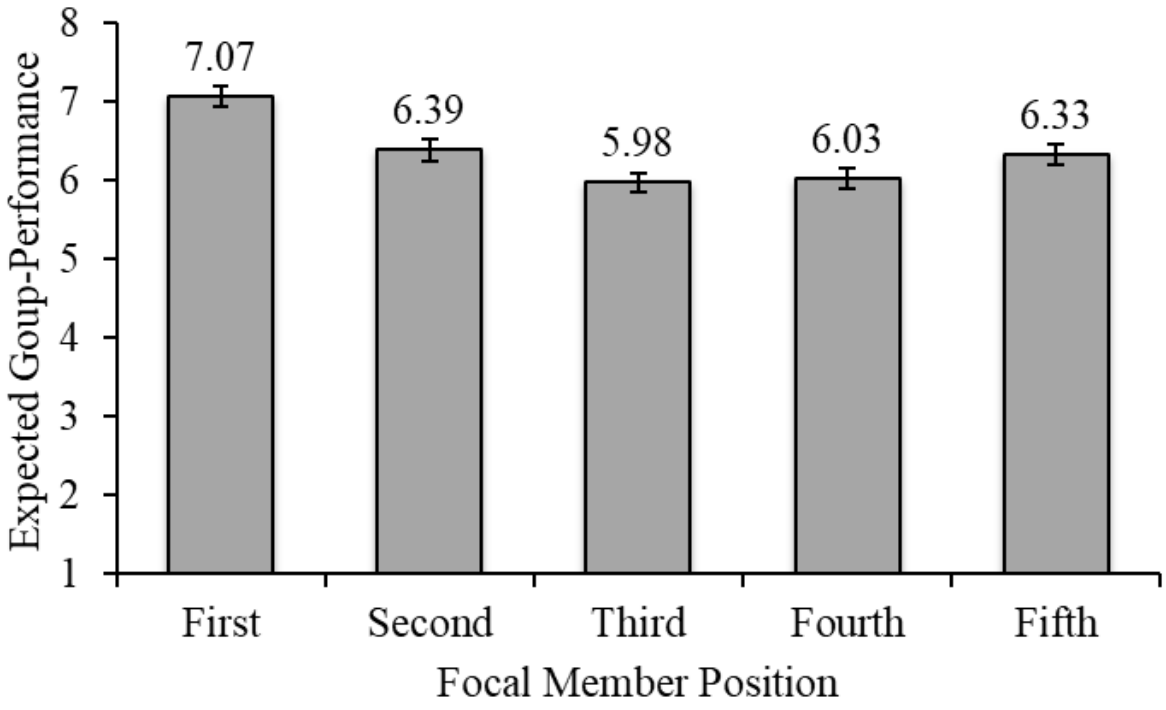

*Figure 1***.** Expected performance of the rest of the group depending on the focal member's position in the sequence in Study 2. The expected group-performance matched that of the first group member more than that of any other member. Error bars represent standard errors.

These results suggest that the first member has a unique status in influencing judgments about the rest of the group. It is not the case that earlier members have greater influence than later members. Rather, the effect is unique to the first. A possible follow up question is whether perceivers assume the rest of the group assimilates to the first member or contrasts from those in other positions. Our hypothesis pertains to the relative greater perceived diagnosticity of the first compared to other group members, but we do not suggest that other group members are non-diagnostic, or that perceivers contrast from them to the group. Indeed, a follow-up study, which included a control condition with no sequence information, showed that people do not contrast

from the other positions, but rather make judgments that assimilate to the first position to a greater degree than to any subsequent position (see Supplemental Material Study 1S). Next, we tested the first-member heuristic in judgments based on poor performance.

Study 3: Good and Bad Performance

In Study 3, we tested whether the first-member heuristic generalizes to poor performance. Participants learned about a runner in a relay race who performed either well or poorly and was either the first, middle, or last runner in his team. They then indicated their expectations about the performance of the other group members. We predicted the performance of the first would have the most impact on judgments of other group members, such that participants would expect the team's performance to be worse or better after learning that the first (vs. middle or last) runner performed poorly or well, respectively.

Method

Participants and Design. We opened the study to 240 participants on MTurk, and 245 participants (104 females; M_{age} = 36.50, SD = 12.06) completed the study in exchange for $0.30. In this study, we recruited approximately 40 participants per condition based on our initial sample-size calculation (see introduction section). The study employed a two-factorial 3 (position of focal group member: first vs. middle vs. last) × 2 (performance of focal group member: good vs. bad) between-subjects design. Six additional participants quit the survey after condition assignment but before completing the dependent variables (two from the first condition and four from the middle condition).

Materials and Procedure. Participants read a scenario about a group of runners in a relay race. As in the previous studies, we emphasized that the positions of the runners were determined randomly (each runner drew his name out of a box) to eliminate inferences about

why the focal member might have been in his particular position in the race. Participants then read about a focal member, John, who was either the first, third, or last to run in the race, and who was either very fast (good performance) or very slow (poor performance). Then, participants indicated how fast they expected each of the other four runners —Mark, Keith, Dan, and Chris—to have run during the race (1 = *very slow*; 9 = *very fast*), α = .93. Before these items, we included two exploratory questions: participants indicated the likelihood that (a) the group would win the race, and (b) that the group would be happy with its performance (1 = *very unlikely*; 9 = *very likely*; see Supplemental Material for these exploratory analyses).

Additionally, to test and control for the possibility that participants have lay theories about the influence that each runner exerts on the group's overall performance (e.g., first runners always have the most influence), we asked them how much of an impact they thought John had on the group performance (1 = *very small impact*; 9 = *very large impact*). If participants believed that all runners have a similar impact on the group's outcome, such lay theories about runners in relay races would not be an alternative explanation for our findings on the first-member heuristic.

At the end of the study, as an attention check, participants indicated John's position in the sequence. Six participants failed the attention check by misreporting the focal member's position (one from the first condition, three from the middle condition, and two from the last condition). We retained these participants in the subsequent analyses and note that excluding them did not change the pattern or the significance of the results.

Results and Discussion

An ANOVA of the focal member's position in the sequence × performance on expected group-performance yielded no main effect of the focal member's performance, $F(1, 239) =$

0.001, $p = .974$, $\eta^2 < 0.01$, and no main effect of his position in the sequence, $F(2, 239) = 0.021$, $p = .980$, $\eta^2 < 0.01$. However, we found the predicted interaction, $F(2, 239) = 6.003$, $p = .003$, $\eta^2 = 0.05$ (see Figure 2).

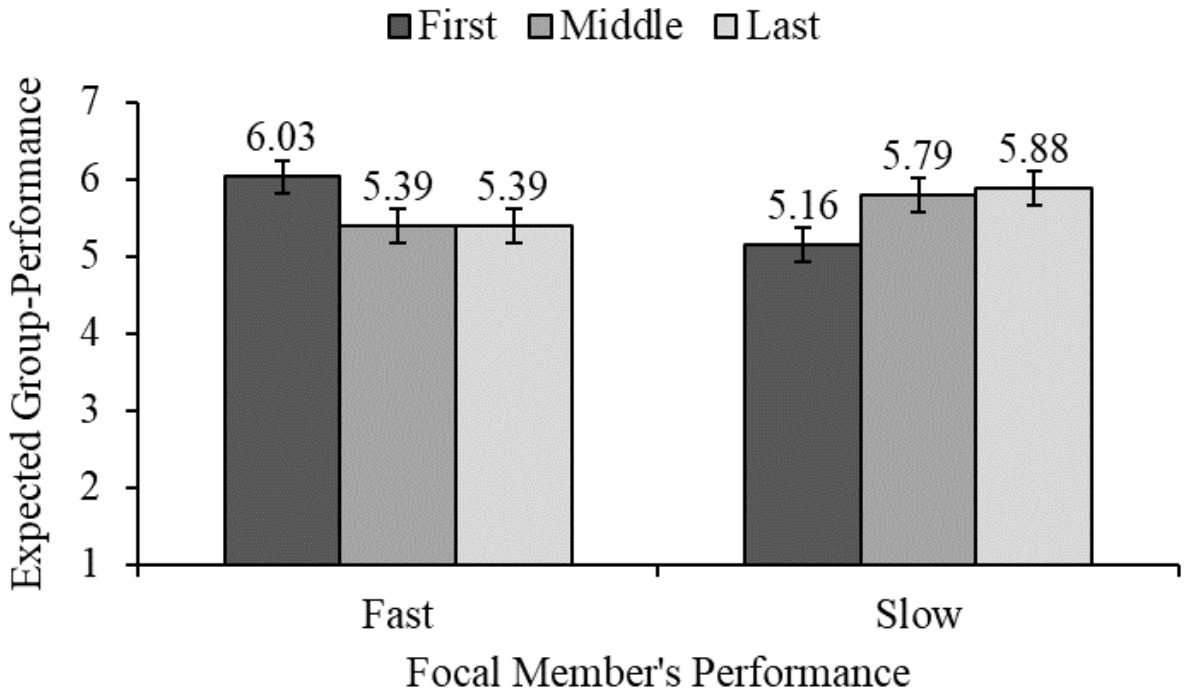

*Figure 2***.** Expected performance of the rest of the group depending on the focal member's performance and position in the sequence in Study 3. The expected group-performance matched that of the first group member more than that of the middle or last group member. Error bars represent standard errors.

Specifically, when the focal member was slow, the rest of the group was expected to be slower if the focal member was in the first position ($M = 5.16$, $SD = 1.58$) relative to the middle ($M = 5.79$, $SD = 1.09$), $p = .042$, CI[-1.245, -0.024], or last position ($M = 5.88$, $SD = 1.32$), $p = .021$, CI[-1.330, -0.109]. There was no difference in expected group-performance when the focal member was in the middle or last position, $p = .783$, CI[-0.696, 0.525]. Additionally, when the focal member was fast, the rest of the group was expected to be faster if the focal member was in

the first position (M = 6.03, SD = 1.12), relative to the middle (M = 5.39, SD = 1.61), p = .043, CI[0.019, 1.256] or last positions (M = 5.39, SD = 1.59), p = .039, CI[0.034, 1.255]. Again, there was no difference in expected group-performance between the middle and the last condition, p = .982, CI[-0.618, 0.604].

Furthermore, as predicted, only the performance of the first (vs. middle or last) member colored participants' expectations about the rest of the group. That is, we did not find a main effect of the focal member's performance, such that a fast (vs. slow) focal member would lead to expectations that the rest of the group was faster, regardless of the runner's position. Rather, the performance of the focal member only affected judgments about the group when that focal member was first (i.e., the other runners were seen as faster if he was fast compared to if he was slow), p = .006, CI[-1.487, -0.258]. However, when the focal member was third, the other runners' predicted performance was not affected by whether he was fast or slow, p = .202, CI[-0.215, 1.013]. Similarly, when the focal member was last, the other runners' predicted performance was not affected by whether he was fast or slow, p = .112, CI[-0.116, 1.098].

Moving to ratings of perceived impact of the focal member's performance on the group's overall performance, we found a main effect of performance, $F(1, 239) = 6.802$, p = .010, η^2 = 0.03, such that participants expected a fast (vs. slow) focal member to have greater impact. There was no main effect of position in the sequence, $F(2, 239) = 0.070$, p = .933, $\eta^2 < 0.01$, and there was no position × performance interaction, $F(2, 239) = 0.791$, p = .455, $\eta^2 < 0.01$. These results suggested that differential perceptions of the impact of an individual group member on the overall group performance are unlikely to account for the first-member heuristic.

The results of Study 3 suggest that the first-member heuristic is not limited to good performances; poor performers also appear more diagnostic for inferences about the rest of the

group when they are first (vs. middle or last). Notably, in the relay context, it is possible that people believe the first member sets the tone for the rest of the group (either inspiring or demoralizing, Antipov & Pokryshevskaya, 2017). We address this possibility in other studies, for example in Studies 4-5, in which the group members perform simultaneously. We next move to a different test of the first-member heuristic by asking whether people seeking to form a judgment of a group would be more likely to examine the performance of the first (vs. any subsequent) member.

Study 4: Looking at the First Member to Judge the Group

In Study 4, we tested whether people trying to form a judgment of a group would be more likely to examine the performance of the first (vs. any subsequent) member. In previous studies, we attempted to minimize inferences based on conversational norms (Grice 1957, 1975) by emphasizing the random or arbitrary nature of the group's order. However, participants might still have made such inferences, assuming that the scenario told them about the first member because it was unfolding sequentially, and about a member in another position because this member was an outlier. Accordingly, in Study 4, we asked participants to select themselves the group member about whom they would like to receive information in order to make a judgment about the entire group. Participants read about a group of research scientists who immigrated to the United States on work visas, and indicated which scientist's performance they would like to review in order to evaluate the entire group. In order to form an accurate assessment of the group, participants should select the group member they expect to be most diagnostic of the entire group. We predicted the highest proportion of them would choose the first member.

Method

Participants and Design. We opened the study to 210 participants on MTurk and 213

participants (99 females; M_{age} = 35.42, *SD* = 12.16) completed the study in exchange for $0.30. We had no a-priori estimate of the required sample size, thus we chose a conservative sample size of 210. The study used no between-participants conditions, that is, all participants received the same information. We pre-registered the data collection and analysis plan for this study at http://aspredicted.org/blind.php?x=wy3kf8.

Materials and Procedure. Participants read a scenario about five cancer researchers who applied for H-1B visas to work in the US and received their visas in an arbitrary order determined by administrative processes related to their birthdays: first Milom, then Jaron, Clovas, Bridi, and finally Erlend. To highlight that the researchers' performances took place simultaneously, the scenario added that all five researcher moved to the US around the same time.

Next, participants read that the president of the university employing the five scientists wanted to assess their performance. Due to time constraints, the president could not review the performance of all five scientists, so he would only select and read the performance review of one scientist to form an impression of the group. We then asked participants to indicate whom they would select if they were the president: 1 = *Milom*, 2 = *Jaron*, 3 = *Clovas*, 4 = *Bridi*, or 5 = *Erlend*.

Finally, as an attention check, we asked participants what had determined the order in which the scientists received their visa (1 = *The best researcher got their visa first*; 2 = *Arbitrary administrative processes having to do with the birthdays determined the order*; 3 = *The researcher whose case was the easiest to process got the visa first*; 4 = *I don't remember*). Sixteen percent of participants falsely reported that the best researcher got their visa first (option "1") and another six percent of participants falsely reported that the researcher whose case was

easiest to process got their visa first (option "3"). These participants thus failed the attention check. Another five percent of participants did not remember (option "4") how the order had been assigned. We retained both the participants who failed the attention check and those who did not remember the order assignment in the subsequent analyses and note that excluding them did not change the pattern or the significance of the results.

Results and Discussion

In line with our prediction, the first member was chosen significantly more often than other members, $\chi 2$ (4) = 119.28, $p < .001$. Of 213 participants, 100 (47%) chose the first group member to learn about the group, whereas 20 (9%) chose the second member, 55 (26%) chose the third member, 17 (8%) chose the fourth member, and 21 (10%) chose the last member.

This study finds that people are most likely to choose to see information about the first member when they want to learn about a group (i.e., the first group member is perceived as the most diagnostic). In a context where conversational norms and influence (i.e., the first member setting the tone) are eliminated, we found support for the first-member heuristic. It appears that people believe the first member in an (arbitrarily) ordered group is more informative for inferences about the rest of the group than any other group member. This pattern of judgment should further translate into the treatment of groups. In the next two studies, we therefore investigated participants' treatment of groups. Specifically, the next study tested whether the performance of the group member in the first (vs. any subsequent) position would have greater influence on participants' support for a policy or decision that would affect the entire group.

Study 5: Supporting Policies against the Group

Moving to the treatment of groups, Study 5 tested whether the poor performance of the group member in the first (vs. middle or last) position would negatively influence judgments of

the group, which in turn would increase support for a policy detrimental to the group. Participants read about a foreign research scientist who performed poorly at his job and had been either the first, middle, or last of the group to receive his work visa to enter the United States. We predicted that participants would expect the rest of the group to perform worse and would therefore be less supportive of extending the group's stay in the country when the focal member was in the first (vs. middle or last) position. We also predicted that expected group-performance would mediate the effect of position on support of the more restrictive immigration decision regarding this group.

Method

Participants and Design. We opened the study to 210 participants on MTurk (80 females; 51% of the sample between 25-34 years old), who completed the study in exchange for $0.85. The sample size in this study was based on our initial sample-size calculation, but we recruited more participants (approximately 70 per condition) to be able to exclude those participants who were not US-citizens (and might thus themselves be affected by immigration policies) while maintaining sufficient statistical power for the analyses. The study employed a one-factorial (position of focal group member: first vs. middle vs. last) between-participants design. Six participants quit the survey after condition assignment but before completing the dependent variables (two from the first condition and four from the middle condition).

Materials and Procedure. Participants read the same scenario as in Study 4 about five cancer researchers who applied for H-1B visas to work in the US. Unlike in Study 4, participants also learned about a focal researcher (Milom), who had been the first (vs. third or last) to obtain his visa to enter the US. They further learned that after several months on his job, he made a grave mistake that jeopardized his lab's work.

Next, to assess judgment of the group, participants indicated how they thought the other researchers—Jaron, Clovas, Bridi, and Erlend—did on their projects (1 = *very poorly*; 9 = *very well*), $\alpha = .89$. To measure treatment of the group, participants read that H1-B visas are initially valid for three years and may be extended for an additional three years, and rated the extent to which they thought the other researchers' visa should be extended (1 = *definitely not*; 9 = *definitely yes*), $\alpha = .93$.

Because immigration is a politically charged topic, we also had participants indicate whether they were eligible to vote in the United States (1 = *yes*; 2 = *no*) to indirectly assess their US citizenship status. Participants also rated their political views (1 = *very liberal*; 9 = *very conservative*) and, as an exploratory measure, they rated their expectations about the degree of connection between Milom and each of the other four researchers (1 = *very weak*; 9 = *very strong*; $\alpha = .88$; see Supplemental Material for the exploratory analyses).

Results and Discussion

In support of the first-member heuristic, participants' predictions of the other group members' performances differed depending on the focal member's position in the sequence, $F(2, 207) = 8.405$, $p < .001$, $\eta^2 = .08$. When the poorly-performing focal member had been the first to receive his visa, expected group-performance was worse ($M = 6.17$, $SD = 1.70$) than when the focal member had been the third ($M = 7.11$, $SD = 1.23$), $p < .001$, CI[-1.428, -0.473], or last to receive his visa ($M = 6.98$, $SD = 1.35$), $p = .002$, CI[-1.304, -0.313]. There was no difference in expected group-performance between the third and the last condition, $p = .534$, CI[-0.308, 0.593].

Furthermore, we found the predicted effect of the focal member's position on participants' support for extending the other researchers' visas, $F(2, 207) = 3.516$, $p = .032$, $\eta^2 =$

.03. When the poorly-performing focal member had been the first to receive his visa, participants were less supportive of extending the other researchers' visas (M = 6.66, SD = 2.10) than when the focal member had been the third (M = 7.38, SD = 1.33), p = .012, CI[-1.270, -0.162], or last to receive his visa (M = 7.27, SD = 1.53), p = .040, CI[-1.176, -0.026]. There was no difference in visa-extension support between the third and the last condition, p = .666, CI[-0.637, 0.408].

A mediation analysis revealed that expected group-performance mediated the effect on visa-extension preferences. Specifically, expected group-performance predicted participants' support for extending the researchers' visas, *Exp(B)* = 0.616, *CI Exp(B)* = [0.486, 0.746], p < .001, R^2 = .295. When controlling for expected group-performance, the effect of the focal member's position on visa-extension preferences became non-significant, *Exp(B)* = 0.050, *CI Exp(B)* = [-0.200, 0.300], p = .691, whereas expected group-performance remained a significant predictor of support for extending the visas, *Exp(B)* = 0.611, *CI Exp(B)* = [0.477, 0.744], p < .001, R^2 = .295. These results showed that expected group-performance mediated the effect of the focal member's position on participants' support for extending the researchers' visas (95% CI Bootstrapping [0.076, 0.449], 1000 bootstrapping samples).[3]

We found that a member's position in an (arbitrarily) ordered group sequence can influence people's opinions on consequential treatment of the rest of the group. Participants in Study 5 expected the other group members to perform worse if the first member (vs. middle or

[3] When excluding participants who were presumably not US-citizens (because they were not eligible to vote, N = 7), the first-member heuristic prevailed on the expected group-performance, F(2, 200) = 9.338, p < .001, η^2 = .09, and on support for extending the visas, F(2, 200) = 3.762, p = .025, η^2 = .04. The pattern and significance of all planned contrasts were the same, whether these participants were included or excluded. Also when controlling for participants' political beliefs, the first-member heuristic on expected group-performance, F(2, 199) = 9.484, p < .001, η^2 = .09, and support for extending the visas remains unchanged, F(2, 199) = 5.237, p = .006, η^2 = .05. Thus, individual differences that might determine participants' attitudes toward immigration did not affect the first-member heuristic for judgment and treatment of the group.

last) had performed poorly, which lead them to express more detrimental policy preferences toward the group. In the next study, we tested whether the performance of the group member in the first (vs. middle or last) position would further have a greater influence on participants' decision to join the group.

Study 6: Choosing to Join a Group

In Study 6, we examined people's decisions to join a group with potential financial implications, based on information about the performance of the first (vs. middle and last) member. Participants chose a group for a game of trivia that offered a bonus payment, based on their group's final performance. They received information about a single player in each of three groups: that player—the first, middle, or last in his/her own group—did slightly worse than the others. We predicted that seeing the poor performer in the first position would deter participants from joining the group to a greater extent than seeing the poor performer in the middle or last position.

Method

Participants and Design. We opened the study to 150 participants on MTurk, and 152 participants (56 females; M_{age} = 34.29, *SD* = 10.23) completed the study on desktop or laptop computers (no mobile phones) in exchange for $0.20. The sample size in this study (approximately 50 participants per condition) was based on our initial sample-size calculation. The study employed a one-factorial (position of displayed group-member: first vs. middle vs. last) between-participants design. In addition to the base payment, participants learned about the opportunity to win a $20 bonus if their group performed well. Because in reality, there were no groups competing against each other for the bonus, we randomly selected one participant at the end of the study to receive the $20 bonus.

Materials and Procedure. Participants read that they would participate in a trivia game and that other MTurk workers from all over the United States were also responding to these questions. They further read that groups of six MTurk workers would form, whose answers would be pooled. The group with the highest overall score would win a prize of $20. Their task was to choose which group they wanted to join. Specifically, they read: *"Because other MTurk workers have already responded to the questions earlier today, we have enough responses to give you some information on the other group members… You have 3 groups to choose from... We will show you the score of one randomly selected group member per group."*

In all three conditions, participants received information about the first player of one group, the third player of another group, and the last player of yet another group (see Figure 3). This information was displayed simultaneously on one screen. Thus, each participant saw information about three players from three different groups. They learned that two of the three players had given eight correct answers, whereas the remaining player had given seven correct answers—a slightly worse performance. We used a small performance difference between players to allow for the possibility that some participants would still select the slightly worse player's group. The only variation between conditions was whether this slightly worse player was the first, third, or last of his/her own group. After receiving information about the three players, participants chose which group they wanted to join. Next, participants were informed that the study was finished and were debriefed.

Note that this experimental design differed from our previous studies, in which participants received information about a single target and predicted the performance of the others (Studies 1-3 and 5) or selected a group member about whom they wanted to receive information (Study 4). In contrast, in this study, participants did not predict the performance of

the other players or choose a player, but instead chose which group to join.

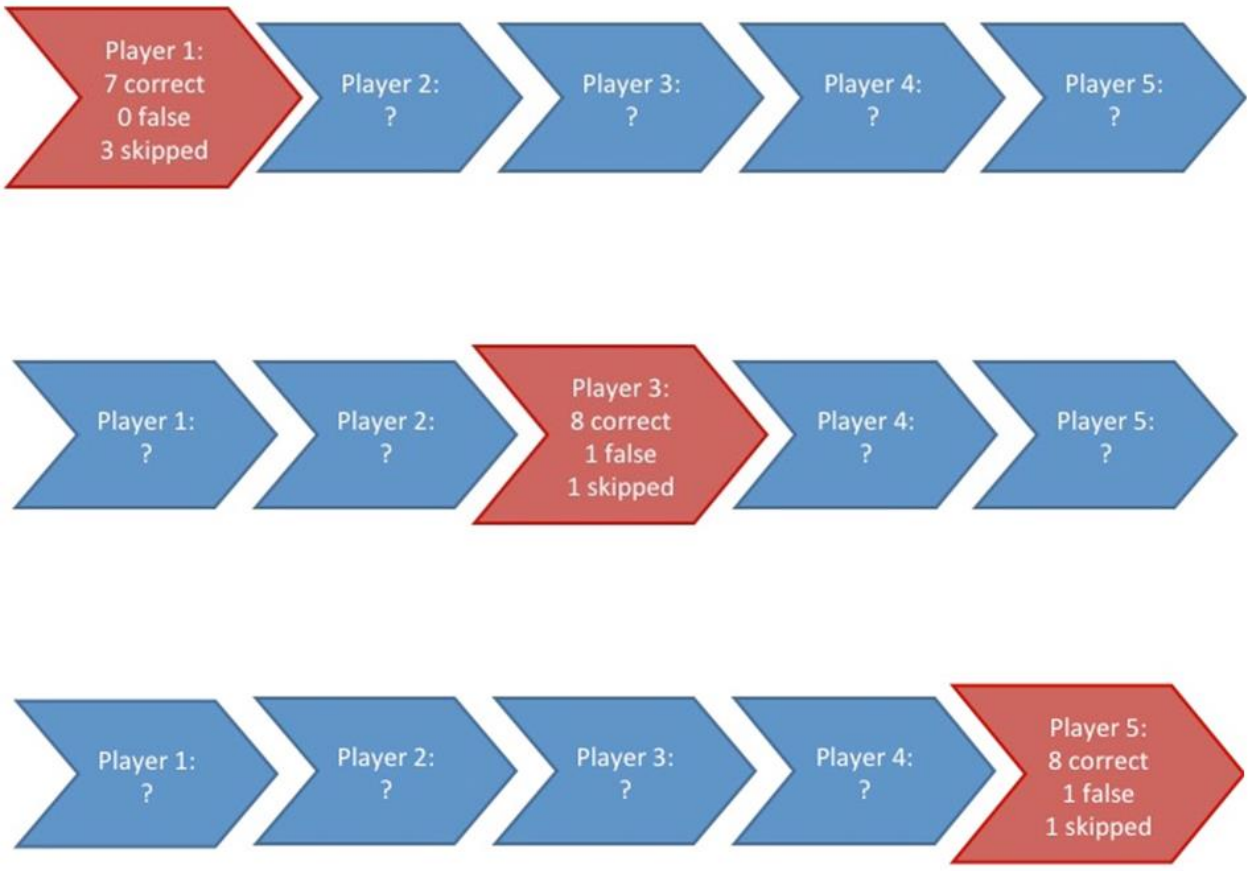

Figure 3. Illustration of the information given to participants in the first condition (i.e., the worst player is in the first position).

Results and Discussion

Participants were overall less likely to choose the group of the worst performer among the three displayed players (i.e., worst displayed player): 26% of participants chose this group, compared to an expected 33% if choices were random. A binominal test revealed that the difference between the observed and the expected proportion was significant, $p = .034$, indicating that participants chose this player's group less frequently, presumably because they recognized that this group includes a relatively low performer.

In support of our hypothesis, an ordinal regression revealed that group choices differed depending on the position of the worst displayed player in his/her own group sequence, $\chi^2(2) = 6.256$, $p = .044$, Nagelkerke's $R^2 = .045$. Whereas only 17% (8/42) of participants chose the group in which the worst displayed player was the first in his/her group, 33% (17/35) chose the group in which the worst displayed player was the third in his/her group's sequence, and 28% (14/36) chose the group in which the worst displayed player was the last in his/her group's sequence. When the worst displayed player was first in his/her group, the odds of participants choosing this player's group were 0.781, CI_{odds}[0.139, 1.423], in comparison to the odds of choosing the worst displayed player's group when this player was third or last of his/her own group, Wald $\chi^2(1) = 5.683$, $p = .017$. There was no difference in choices between the conditions in which the worst displayed player was third or last of his/her group, Wald $\chi^2(1) = 0.385$, $p = .535$. We note that, if participants had chosen randomly, the expected odds for each choice would be 1.000, which would correspond to 33.33% of participants choosing each of the three options. However, we found that the odds of choosing the group of the worst performer among the three displayed players were lower than would be expected from random chance (0.781).

Thus, in support of our hypothesis about the first-member heuristic, the performance of the first player influenced participants' choices more than the performance of the third or last player. In our studies so far, we have documented the first-member heuristic in the judgment and treatment of a wide variety of groups (e.g., gymnasts, runners, foreign scientists, trivia contestants). Although we posit and show that the first-member heuristic occurs in the evaluation of groups, it is still unclear whether it might also extend to the evaluation of any collection of ordered items (e.g., five random strangers ordered from 1 to 5) or to broad categories of items (e.g., all gymnasts). In our final study, we tested whether the perceived homogeneity of a

group—that is, the perception that a collection of members are similar to each other and form a group—is a necessary precondition for this effect. Additionally, our next study moved to groups of racehorses to test whether the effect extends beyond the evaluation of groups of people.

Study 7: Moderation by Group Homogeneity

In Study 7, we tested whether group homogeneity moderates the first-member heuristic. We predicted that the first-member heuristic occurs for the evaluation of groups, but not for the evaluation of any ordered collection of items or broad category. We operationalized homogeneity through similarity along a salient dimension. Participants read about five racehorses who each took part in their own separate races on one day. Depending on the condition, the five racehorses where described as having trained together with a team of closely related trainers (high-homogeneity condition) or as having trainers who had never met each other (low-homogeneity condition). Furthermore, the focal racehorse performed poorly and was either the first, middle, or last of the five to race that day. Participants then indicated their expectations about the performance of the other racehorses. We predicted the first-member heuristic would occur in the high-homogeneity condition, but not in the low-homogeneity condition. Indeed, a collection of items (e.g., racehorses) that is low in homogeneity should be less likely to be considered a group, and each member should be seen as uninformative for inferences about the rest of the group.

Method

Participants and Design. We recruited 300 student participants from different campus locations of a Dutch university (187 females; M_{age} = 22.23, SD = 5.19) in exchange for a small candy bar. The study employed a 3 (position of focal group member: first vs. middle vs. last) × 2 (group homogeneity: high vs. low) between-participants design.

Materials and Procedure. Participants read a scenario about five racehorses, each racing in a different category on the same day. In the high-homogeneity condition, a team of closely related trainers had trained the horses together, whereas in the low-homogeneity condition, the trainers had never met each other and the horses had been trained apart (for a similar manipulation, see Hamilton et al., 2015). Participants further read that the focal horse Flying Mane, who had run either first, third, or last, had been slow and had lost the race. Next, participants rated how they expected the other horses' races went—Storm Lion, Red Runner, Silver Dream, and Sandy Bird (1 = *very bad*; 9 = *very good*), α = .50.

Results and Discussion

An ANOVA of the focal horse's position in the sequence × group homogeneity on expected group-performance yielded a main effect of position, $F(2, 293) = 5.267$, $p = .010$, $\eta^2 = .03$. Participants expected the four horses to have run slower when the focal horse was first, compared to middle or last. Furthermore, we found an unexpected marginally significant main effect of group homogeneity, $F(1, 293) = 3.483$, $p = .063$, $\eta^2 = .01$, such that the four horses were perceived as having run slightly faster when the trainers had been a team (high homogeneity).

Importantly, in support of our hypothesis, we found the predicted interaction, $F(2, 293) = 3.925$, $p = .021$, $\eta^2 = .03$ (see Figure 4). In the high homogeneity condition, we replicated the first-member heuristic: When the focal racehorse ran first on the day of the race and performed poorly, participants expected the other horses to be slower in their separate races ($M = 5.72$, $SD = 1.05$) than when the focal horse ran third ($M = 6.20$, $SD = 0.88$), $p = .025$, CI[-0.904, -0.062] or last that day ($M = 6.57$, $SD = 1.11$), $p < .001$, CI[-1.272, -0.438]. Unexpectedly, the difference between the middle and last condition was marginally significant, $p = .083$, CI[-0.783, 0.049].

Because we found no such difference between the middle and last conditions in other studies, we attribute this finding to chance. In the low homogeneity condition, we found no difference in expected group-performance when the focal horse was first (M = 5.87, SD = 1.34), middle (M = 6.04, SD = 0.90), or last to run on the day of the race (M = 5.89, SD = 1.01), all ps > .409.

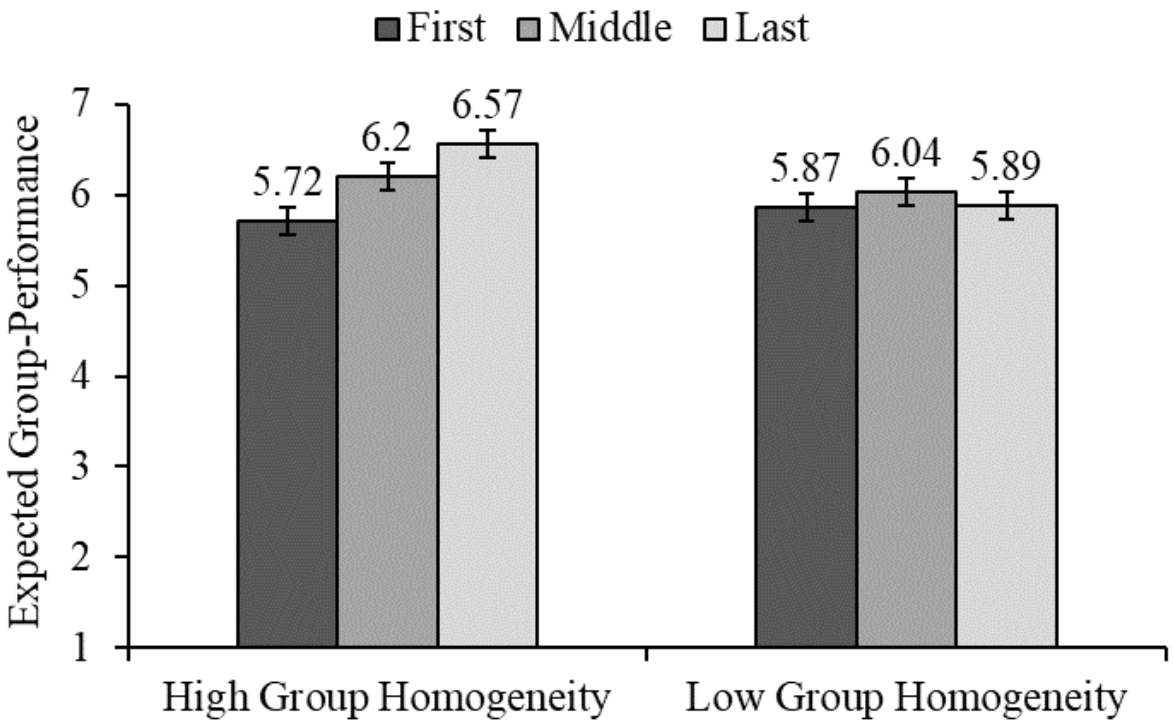

*Figure 4***.** Expected performance of the rest of the group depending on the focal member's position in the sequence and on the group's homogeneity in Study 7. Only for homogenous groups, the expected group-performance matched that of the first member more than that of the middle or last member. Error bars represent standard errors.

Study 7 confirmed our hypothesis regarding the moderating role of group homogeneity on the first-member heuristic. We found that the group member in the first position influenced expectations about the rest of the group when the group was homogenous, but not when the group was heterogeneous. Finally, the study also showed that the first-member heuristic is not limited to the evaluation of groups of people, but also applies to groups of animals.

General Discussion

Across seven studies, we found that people perceive the performance of the first member in an arbitrarily ordered group to be more diagnostic for inferences about the group's performance than the performance of subsequent members (e.g., middle or last). The performance of other members matches that of the group member in the first position more closely than that of group members in other positions. Differences in expected group-performance in turn influence people's treatment of the group. We termed this effect the "first-member heuristic."

Specifically, we found that participants drew stronger inferences from the performance of the first (vs. any subsequent) group member about the performance of the rest of the group (Studies 1-2). This pattern of response emerged regardless of whether the first (vs. middle or last) member performed well or poorly (Study 3). Furthermore, when trying to evaluate or predict a group's performance, participants were more likely to examine the performance of the first (vs. any subsequent) member of the group (Study 4). Because first members are deemed more informative for people's predictions about the group's performance than other members, first (vs. middle or last) members had a greater impact on people's treatments of groups. Participants were less willing to endorse policies that would benefit the group if the first (vs. middle or last) member performed poorly because they then expected a lower performance from others in the group (Study 5). When choosing a group for the chance to win a performance-based monetary reward, participants avoided groups in which the first (vs. middle or last) member had performed relatively poorly (Study 6). Finally, we found that the perceived homogeneity of the group moderated the first-member heuristic effect, such that it did not occur for nonhomogeneous group (Study 7). This last result indicated that the first-member heuristic

applies to the judgment and treatment of groups (collection of similar items) rather than to any collection of items or broad categories.

Previous research has identified circumstances under which first members might dominate judgments about other members or about the group as a whole (e.g., better memory for first members, or comparison of other members to first members). Recognizing the importance of order effects on judgments, prior research has investigated the evaluation of first versus later positions in sports or music competitions (Antipov & Pokryshevskaya, 2017; Damisch, Mussweiler, & Plessner, 2006; de Bruin, 2006). Our work extends this research by investigating whether first members are diagnostic of other members independent of memory or comparison effects. To do so, in Studies 1-3, 5 and 7, participants learned how one particular group member (e.g., first/middle/last) had performed in a race, a project, or a gymnastic routine. Then, we asked participants how they expected the other group members to have performed. With this design, we presented the critical information about one group member to participants without any opportunity to memorize or to compare information about the first member to information about other members. Thereby, we showed that the first-member heuristic is unlikely to originate from memory or comparison effects.

Furthermore, we explicitly told our participants that the first members were arbitrarily determined to be first (Studies 1-4, and Studies 1S and 2S in the Supplemental Material). Thereby, we tested the first-member heuristic in the absence of specific circumstances that might have rendered the first member special (e.g., first-is-best assumption, Carney & Banaji, 2012). We also ensured by experimental design that participants could not have inferred that the first somehow influenced or foreshadowed other group members (Studies 4-6). Furthermore, we addressed conversational norms as an alternative explanation by telling participants that the

computer would select on which group member they saw information (Studies 1 and 6) or by asking them to select a group member (Study 4). Across these studies, the first member consistently affected judgments and treatments of the rest of the group to larger extents.

Theoretical implications

These findings have several theoretical implications. First, they contribute to the literature on diagnosticity (Hamilton et al., 2015), as we found that group members can vary in their diagnosticity. Whereas past research found that diagnosticity depends on the group's similarity, we found that within a given (homogenous) group, first members are more diagnostic than other members.

Second, these findings contribute to the literature on judgments based on sequential positions, which has identified several reasons why first positions might affect judgments more than other positions (e.g., comparisons, foreshadowing; Antipov & Pokryshevskaya, 2017; LeBoeuf et al., 2014). We find that first members can especially impact judgment and treatment of the group even in the absence of these reasons.

Third, these findings contribute to the study of heuristics. We argued that people overgeneralize from contexts in which first members are actually more diagnostic than other members (e.g., first employees shaping a company's workforce). Thereby, the first-member heuristic follows the same logic as other social heuristics, for example, the heuristic to cooperate even in one-shot anonymous games (Rand et al., 2014). Research has shown that such heuristics are overly applied to judgment and decision-making but has also shown that the application of such heuristics can be reduced through ample decision time and experience with tasks such as one-shot games (Rand et al., 2014; Yamagishi et al., 2007).

The question arises whether ample decision time and experience would also lead people

to apply the first-member heuristic to lesser extents. We think that increasing the time that people have to make a judgment about a group and decide how to treat the group in itself would not reduce the use of the first-member heuristic. In Study 6, people had an incentive to select the best group supposedly to maximize their chances to win the bonus, and they had no time pressure to select the group. Thus, people presumably took the time they needed to make their decision, but despite the motivation and time to select the best-performing group, the first-member heuristic nevertheless emerged. Regarding experience, the repeated exposure to a group whose first member is not more diagnostic than other members should decrease people's use of the first-member heuristic. However, this learned limitation of the rule that first members are more diagnostic would presumably only be used for the specific experimental contexts in which it is learned. Parallel to other social heuristics, experience with one-shot games decreases cooperation in these games but presumably not in people's life overall.

One might wonder to what types of judgments people apply the first-member heuristic. We conducted a supplemental study testing a boundary condition of the first-member heuristic (see Supplemental Study 2S): We replicated the first-member heuristic for information that is relevant to group membership (i.e., a member's professional performance for a group of professionals). However, we showed that the first-member heuristic disappears for information unrelated to the group context (i.e., a member of a group of professionals being kind-hearted and generous). This supplemental study is in line with our overgeneralization explanation: First group members that truly influence the selection of subsequent group members will most likely influence this selection on characteristics and traits related to the group context.

More generally, we expect the first-member heuristic to occur for qualities that are likely similar between group members. For example, gymnasts from the same team or researchers in

the same university likely have similar abilities, because their abilities are relevant to becoming a group member. If gymnasts had vastly different abilities, they would likely not be part of the same team. Note that we do not expect collaboration between group members to be a necessary condition for the first-member heuristic, as competitors in the same contest also likely have similar abilities. In such instances, we would expect the first competitor to be more diagnostic than other competitors. Thus, we expect the first-member heuristic for qualities that are relevant to the group context (e.g., abilities, knowledge, competence in a group of scientists) but not for qualities that are irrelevant to the group context (e.g., generosity, kind-heartedness in a group of scientists).

With regard to the implications of the first-member heuristic for research on the prototypicality of group members, one could speculate that a particularly diagnostic group member (e.g., the first member) might be seen as prototypical for the group, that is, as having the characteristics that are normative for belonging to the group (Hogg, Hains, & Mason, 1998). Such a reasoning might follow from the same overgeneralization: If the first member was instrumental in selecting and shaping subsequent members, the standards for a normative and prototypical member might also have been shaped by the first member. The more prototypical a given group member, the more that member seems to fit the requirements of being the group leader, even if that member does not conform to a leader stereotype (Hogg et al., 1998). Thus, prototypicality overwrites leadership stereotypes when people evaluate potential group leaders. If first members appear more diagnostic, they might also appear more prototypical for the group and might thus be more likely to be selected as group leaders.

Practical implications

The first-member heuristic also has several important practical implications. First, we

found that people treat groups differently based on information on the first (vs. other) member. For example, people's support of policies against the group (Study 5) depends on whether an offender that is a group member was first versus not. Second, in daily life, people, animals or objects are often ordered into arbitrary sequences (e.g., student presentation number 1 or reviewer 1). Our studies show that first members are diagnostic of other members even when observers are fully aware that the sequence is arbitrary and meaningless. Thus, groups would be wise to place a strong group member in the first position to raise people's expectations about the rest of the group even when sequences are clearly arbitrary. Similarly, stores or brands should take special care to make customers' experiences with first service providers (cashier number 1) or products (item number 1 of a series, meal number 1 of a menu) especially positive.

Third, our results suggest that a simple reminder of the arbitrariness of a sequence is not sufficient to eliminate sequence effects on judgments. When performances take place in sequences (e.g., gymnastic performances, song contests), the sequence of performance is often randomly determined (e.g., by drawing numbers). Thereby, sequential inferences (e.g., the first is best, Carney & Banaji, 2012) are presumably ruled out. Yet, our studies show that people might still hold beliefs that the sequential position of a person carries special meaning, despite having been informed otherwise. Our results further suggest that one way to eliminate the first-member heuristic would be to draw people's attention to factors that make the group seem less homogeneous.

Limitations

One potential limitation of the present research is that in a subset of our studies, we did not explicitly tell participants that the sequence was arbitrarily determined (Studies 5-7). For example, in Study 6, participants might have inferred that the first person to answer the trivia

questions was the most knowledgeable. This reasoning would provide an alternative explanation for the results. However, participants might have also followed the opposite reasoning, namely that the first person to answer the trivia questions knew so few answers that he/she worked on the quiz the shortest. Although some lay theories of participants might have affected the results of this study, we demonstrate the first-member heuristic also when the sequence is completely arbitrarily assigned (Studies 1-4, and Supplemental Studies 1S and 2S).

Another potential limitation is that our studies used relatively small groups of five or seven members. A follow up question is whether the first-member heuristic would affect judgment and treatment of larger groups. In such large groups—for example, immigrating scientists, gymnast on an athletic team—the definition of "first" might also expand. It could include for example, the first cohort, or the first wave of immigrants.

Taken together, our findings on the first-member heuristic show the important role of arbitrary positions in a sequence for judgments and behaviors related to groups. Whereas previous research has demonstrated that people are more likely to draw inferences about groups from diagnostic (versus less diagnostic) others, we demonstrate that being first in a sequence leads a particular group member to appear more diagnostic of the group than any other member.

References

Anderson, N. H. (1973). Serial position curves in impression formation. *Journal of Experimental Psychology, 97*, 8-12.

Antipov, E. A., & Pokryshevskaya, E. B. (2017). Order effects in the results of song contests: Evidence from the Eurovision and the New Wave. *Judgment and Decision Making, 12*, 415-419.

Bar-Hillel, M. (2015). Position effects in choice from simultaneous displays: A conundrum solved. *Perspectives on Psychological Science, 10*, 419-433.

Birnbaum, M. H. (1972). Morality judgments: Tests of an averaging model. *Journal of Experimental Psychology, 93*, 35-42.

Campbell, D. T. (1958). Common fate, similarity, and other indices of the status of aggregates of persons as social entities. *Behavioral Science, 3*, 14-25.

Carney, D. R., & Banaji, M. R. (2012). First is best. *PloS one, 7*, e35088.

Cone, J., & Ferguson, M. J. (2015). He did what? The role of diagnosticity in revising implicit evaluations. *Journal of Personality and Social Psychology, 108*, 37-57.

Crawford, M. T., Sherman, S. J., & Hamilton, D. L. (2002). Perceived entitativity, stereotype formation, and the interchangeability of group members. *Journal of Personality and Social Psychology, 83*, 1076-1094.

Dai, H., Milkman, K. L., & Riis, J. (2014). The fresh start effect: Temporal landmarks motivate aspirational behavior. *Management Science, 60*, 2563-2582.

Damisch, L., Mussweiler, T., & Plessner, H. (2006). Olympic medals as fruits of comparison? Assimilation and contrast in sequential performance judgments. *Journal of Experimental Psychology: Applied, 12*, 166-178.

de Bruin, W. B. (2006). Save the last dance II: Unwanted serial position effects in figure skating judgments. *Acta Psychologica, 123*, 299-311.

Einhorn, H. J., & Hogarth, R. M. (1986). Judging probable cause. *Psychological Bulletin, 99*, 3-19.

Gilovich, T. (1981). Seeing the past in the present: The effect of associations to familiar events on judgments and decisions. *Journal of Personality and Social Psychology, 40*, 797-808.

Glanzer, M., & Cunitz, A. R. (1966). Two storage mechanisms in free recall. *Journal of Verbal Learning and Verbal Behavior, 5*, 351-360.

Grice, H. P. (1957). Meaning. *The Philosophical Review, 66*, 377-388.

Grice, H. P. (1975). Logic and conversation. In P. Cole & J. L. Morgan (Eds.), *Syntax and semantics: Vol. 3. Speech acts* (pp. 41-58). New York: Academic Press.

Hamilton, D. L., Chen, J. M., Ko, D. M., Winczewski, L., Banerji, I., & Thurston, J. A. (2015). Sowing the seeds of stereotypes: Spontaneous inferences about groups. *Journal of Personality and Social Psychology, 109*, 569-588.

Hamilton, D. L., Sherman, S. J., & Castelli, L. (2002). A group by any other name—The role of entitativity in group perception. *European Review of Social Psychology, 12*, 139-166.

Hogg, M. A., Hains, S. C., & Mason, I. (1998). Identification and leadership in small groups: Salience, frame of reference, and leader stereotypicality effects on leader evaluations. *Journal of Personality and Social Psychology, 75*, 1248-1263.

Holyoak, K. J., & Simon, D. (1999). Bidirectional reasoning in decision making by constraint satisfaction. *Journal of Experimental Psychology: General, 128*, 3-31.

LeBoeuf, R. A., Williams, E. F., & Brenner, L. A. (2014). Forceful phantom firsts: Framing experiences as firsts amplifies their influence on judgment. *Journal of Marketing*

Research, 51, 420-432.

Lickel, B., Hamilton, D. L., Wieczorkowska, G., Lewis, A., Sherman, S. J., & Uhles, A. N. (2000). Varieties of groups and the perception of group entitativity. *Journal of Personality and Social Psychology, 78,* 223-246.

MacFie, H. J., Bratchell, N., Greenhoff, K., & Vallis, L. V. (1989). Designs to balance the effect of order of presentation and first-order carry-over effects in hall tests. *Journal of Sensory Studies, 4,* 129-148.

Mantonakis, A., Rodero, P., Lesschaeve, I., & Hastie, R. (2009). Order in choice: Effects of serial position on preferences. *Psychological Science, 20,* 1309-1312.

McPherson, M., Smith-Lovin, L., & Cook, J. M. (2001). Birds of a feather: Homophily in social networks. *Annual Review of Sociology, 27,* 415-444.

Menon, G., Raghubir, P., & Schwarz, N. (1995). Behavioral frequency judgments: An accessibility-diagnosticity framework. *Journal of Consumer Research, 22,* 212-228.

Murdock, B. B., Jr. (1962). The serial position effect of free recall. *Journal of Experimental Psychology, 64,* 482-488.

Nisbett, R. E., Zukier, H., & Lemley, R. E. (1981). The dilution effect: Nondiagnostic information weakens the implications of diagnostic information. *Cognitive Psychology, 13,* 248-277.

Page, L., & Page, K. (2010). Last shall be first: A field study of biases in sequential performance evaluation on the Idol series. *Journal of Economic Behavior & Organization, 73,* 186-198.

Pandelaere, M., Millet, K., & Van den Bergh, B. (2010). Madonna or Don McLean? The effect of order of exposure on relative liking. *Journal of Consumer Psychology, 20,* 442-451.

Peetz, J., & Wilson, A. E. (2013). The post-birthday world: Consequences of temporal landmarks for temporal self-appraisal and motivation. *Journal of Personality and Social Psychology, 104*, 249-266.

Rand, D. G., Peysakhovich, A., Kraft-Todd, G. T., Newman, G. E., Wurzbacher, O., Nowak, M. A., & Greene, J. D. (2014). Social heuristics shape intuitive cooperation. *Nature Communications, 5*, 3677.

Reeder, G. D. (1985). Implicit relations between dispositions and behaviors: Effects on dispositional attribution. In J. H. Harvey & W. Gifford (Eds.), *Attribution: Basic Issues and Applications* (p. 87-116), New York: Academic Press.

Reeder, G. D., & Brewer, M. B. (1979). A schematic model of dispositional attribution in interpersonal perception. *Psychological Review, 86*, 61-79.

Risen, J. L., Gilovich, T., & Dunning, D. (2007). One-shot illusory correlations and stereotype formation. *Personality and Social Psychology Bulletin, 33*, 1492-1502.

Shteingart, H., Neiman, T., & Loewenstein, Y. (2013). The role of first impression in operant learning. *Journal of Experimental Psychology: General, 142*, 476-488.

Skowronski, J. J. (2002). Honesty and intelligence judgments of individuals and groups: The effects of entity-related behavior diagnosticity and implicit theories. *Social Cognition, 20*, 136-169.

Skowronski, J. J., & Carlston, D. E. (1987). Social judgment and social memory: The role of cue diagnosticity in negativity, positivity, and extremity biases. *Journal of Personality and Social Psychology, 52*, 689-699.

Smith, E. R., & Zarate, M. A. (1992). Exemplar-based model of social judgment. *Psychological Review, 99*, 3-21.

Spencer-Rodgers, J., Williams, M. J., Hamilton, D. L., Peng, K., & Wang, L. (2007). Culture and group perception: Dispositional and stereotypic inferences about novel and national groups. *Journal of Personality and Social Psychology, 93,* 525-543.

Sullivan, J. (2019). The primacy effect in impression formation: Some replications and extensions. *Social Psychological and Personality Science, 10,* 432-439.

Sulmont-Rossé, C., Chabanet, C., Issanchou, S., & Köster, E. P. (2008). Impact of the arousal potential of uncommon drinks on the repeated exposure effect. *Food Quality and Preference, 19*, 412-420.

Susskind, J., Maurer, K., Thakkar, V., Hamilton, D. L., & Sherman, J. W. (1999). Perceiving individuals and groups: Expectancies, dispositional inferences, and causal attributions. *Journal of Personality and Social Psychology, 76*, 181-191.

White, G. L., & Shapiro, D. (1987). Don't I know you? Antecedents and social consequences of perceived familiarity. *Journal of Experimental Social Psychology, 23,* 75-92.

Wyer, R. S., & Hinkle, R. L. (1976). Informational factors underlying inferences about hypothetical persons. *Journal of Personality and Social Psychology,* 34, 481-495.

Yamagishi, T., Terai, S., Kiyonari, T., Mifune, N., & Kanazawa, S. (2007). The social exchange heuristic: Managing errors in social exchange. *Rationality and Society, 19,* 259-291.

Zhao, M., Lee, L., & Soman, D. (2012). Crossing the virtual boundary: The effect of task-irrelevant environmental cues on task implementation. *Psychological Science, 23*, 1200-1207.

Supplemental Material

This supplement contains:

Exploratory analyses in Studies 3 and 5

Supplemental Studies 1S and 2S

Exploratory analyses in Study 3

Study 3 included an exploratory variable of predicted victory—i.e., whether participants thought the group would win. We found a main effect of performance, $F(1, 239) = 40.431$, $p < .001$, $\eta^2 = 0.15$, as the group-victory was seen as more likely when the focal member was fast. There was no main effect of the focal member's position, $F(2, 239) = 0.621$, $p = .538$, $\eta^2 = 0.01$. There was also no significant interaction effect of the focal member's position and his performance on group-victory, $F(2, 239) = 1.529$, $p = .219$, $\eta^2 = 0.01$.

Specifically, when the focal member was first, the group-victory was seen as more likely if he was fast ($M = 5.35$, $SD = 1.78$) compared to if he was slow ($M = 3.17$, $SD = 1.64$), $p < .001$, CI[-3.038, -1.321]. Similarly, when the focal member was third, the group-victory was perceived as more likely he was fast ($M = 4.93$, $SD = 2.12$) than when he was slow ($M = 3.80$, $SD = 1.81$), $p = .011$, CI[-1.978, -0.262]. The same pattern emerged when the focal member was last, as the group-victory was predicted to be more likely when he was fast ($M = 5.33$, $SD = 2.32$) than when he was slow ($M = 3.85$, $SD = 2.03$), $p = .001$, CI[-2.328, -0.632].

This study also included an exploratory variable of group-happiness—i.e., whether participants thought the group would be happy with their results. We found a main effect of performance, $F(1, 239) = 66.164$, $p < .001$, $\eta^2 = 0.22$, as the group was predicted to be happier when the focal member was fast. There was no main effect of the focal member's position, $F(2, 239) = 1.719$, $p = .181$, $\eta^2 = 0.01$. There was also no significant interaction effect of the focal member's position and his performance on group-happiness, $F(2, 239) = 2.378$, $p = .095$, $\eta^2 = 0.02$.

Specifically, when the focal member was first, group-happiness was seen as greater if he was fast ($M = 6.53$, $SD = 1.52$) compared to if he was slow ($M = 4.05$, $SD = 1.80$), $p < .001$, CI[-

3.258, -1.695]. Similarly, when the focal member was third, the group was perceived as happier he was fast (M = 6.38, SD = 1.46) than when he was slow (M = 5.12, SD = 2.04), p = .002, CI[-2.035, -0.471]. The same pattern emerged when the focal member was last, as the group was predicted to be happier when he was fast (M = 6.64, SD = 2.03) than when he was slow (M = 4.80, SD = 1.75), p < .001, CI[-2.610, -1.066].

In this study, although we found the expected interaction of the focal member's position in the sequence and his performance on our dependent variable of expected group-performance, we found no such interaction on expected group happiness and expected group victory. We believe that the lack of effect on these two items is specific to this scenario. We explicitly told participants that the runners are amateurs who have no experience with relay races. We did this do add credibility to the random assignment of runners. Yet, this information likely undermined participants' belief that the group could win or would likely be happy with their result, regardless of the sequential position of the focal member.

Exploratory analyses in Study 5

We did not find that the first was more strongly connected to the rest of the group than the middle or last, $F(2, 205) = 2.798$, p = .063, η^2 = .03. When the poorly-performing focal member's visa was approved first (M = 4.60, SD = 1.82), the focal member was seen as somewhat less strongly connected to the other members than when the focal member was middle (M = 5.16, SD = 1.69), p = .056, CI[-1.140, 0.014]. In the middle condition, the focal group member was seen as more strongly connected to the group than in the last condition (M = 4.59, SD = 1.57), p = .039, CI[0.030, 1.119]. There was no difference between the middle and the last condition, p = .969, CI[-0.584, 0.607].

These results indicate that participants correctly recognize that within the scenario, no

reason existed for the first to be diagnostic because of a special connection to the other members. Thus, all group members are similarly connected to each other. When all group members share similar levels of connection, the first is nevertheless more diagnostic.

Supplemental Study 1S: Adding a Control Condition

Study 1S replicated the first-member heuristic and additionally used a control condition without any sequential information to explore whether the first is more diagnostic or other members are less diagnostic compared to members about whom no further information is known. We predicted that the first group member would be more diagnostic than the middle/last members, and that the member in the control condition would fall in between the first member and the middle/last members.

Method

Participants and Design. We opened the study to 400 participants on MTurk and 406 participants (199 females; M_{age} = 36.94, *SD* = 12.45) completed the study on desktop or laptop computers (no mobile phones) in exchange for $0.30. Because Study 1S replicated the scenario used in Studies 1 and 2, we used the same sample size as in these studies (100 participants per condition). The study employed a one-factorial (position of focal group member: first vs. middle vs. last vs. control) between-subjects design. Another six participants quit the survey after condition assignment but before completing the dependent variables (two from the first condition, one from the middle condition, two from the last condition, and one from the control condition). We pre-registered our data collection and analysis plan at http://aspredicted.org/blind.php?x=u8yv5k.

Materials and Procedure. Participants in the first, middle, or last condition read the same scenario about a group of five gymnasts as in Studies 1 and 2. In the control condition,

participants learned that one of the gymnasts performed her floor routine flawlessly and gave a stunning performance, without any sequential information.

Next, to assess the effect of the focal gymnast's performance on expectations about the other gymnasts' performances, we asked participants on a single item how they thought the rest of the group performed (1 = *very poorly*; 9 = *very well*). We used this single item because in the control condition we did not identify any member by name, thus it would be confusing for participants to be asked about the other gymnasts by name. Finally, participants answered the same attention check as in Studies 1 and 2. Control participants did not answer this item because the drawing was not part of the control scenario. Eleven percent of participants falsely reported that the best gymnast had gone first and thus failed the attention check (12 from the first condition, five from the middle condition, and one from the last condition). Another eight percent of participants reported that they did not remember how the order had been assigned (five from the first condition, seven from the middle condition, and 12 from the last condition). We retained both the participants who failed the attention check and those who did not remember the order assignment in the subsequent analyses and note that excluding them did not change the pattern or the significance of the results.

Results and Discussion

In support of the hypothesis about the first-member heuristic, expected group-performance varied based on the focal member's position, $F(3, 403) = 4.806$, $p = .003$, $\eta^2 = .04$. Specifically, when the good performer was in the first position, participants expected the rest of the group to have performed better ($M = 7.37$, $SD = 1.34$) than when the good performer was middle ($M = 6.75$, $SD = 1.22$), $p = .001$, CI[0.269, 0.967], or last ($M = 6.97$, $SD = 1.20$), $p = .024$, CI[0.052, 0.752]. There was no difference between the middle and last condition, $p = .227$, CI[-

0.135, 0.565]. The control condition was between the first and the last (M = 7.24, SD = 1.26). It did not differ significantly from the first condition, p = .440, CI[-0.486, 0.212] or from the last condition, p = .137, CI[-0.085, 0.615]. Unexpectedly, however, if the performer in the control condition did well, participants expected the rest of the group to have performed better than if the middle performer did well, p = .007, CI[0.131, 0.829].

In the absence of any sequence information, participants perceived the performer as slightly less diagnostic than the first member and as slightly more diagnostic than the last member. Presumably, both "more" and "less" refer to a comparison to the other group members. Thus, members are more or less diagnostic within the comparison frame of their group, not in absolute terms.

Supplemental Study 2S: Type of Information as a Boundary Condition

Study 2S tested a potential boundary condition of the first-member heuristic, namely what type of information participants learned about the first (vs. middle/last) member. We expected that the first (vs. middle/last) member would be more diagnostic when the information about that focal member reflects qualities that are likely relevant to group membership. For example, being a member of a group of professionals requires each member to have similar training and abilities that are relevant for the profession. Thus, if one member performs well, this can indicate that other members are also likely to perform well. In contrast, being a member of a group of professionals is unlikely to require each member to share qualities that are unrelated to the profession (e.g., being generous). Thus, if one member is generous, this does not indicate that other members are also generous. We therefore expected to replicate the first-member heuristic when participants receive information that is relevant for the focal member's group membership, but not when they receive information that is irrelevant for the focal member's group

membership.

Method

Participants and Design. We opened the study to 600 participants on MTurk (324 females; M_{age} = 36.47, *SD* = 12.21), who completed the study on desktop or laptop computers (no mobile phones) in exchange for $0.30. Consistent with Supplemental Study 1S, we used the same sample size of 100 participants per condition. The study employed a 3 (position of focal group member: first vs. middle vs. last) × 2 (information type: relevant vs. irrelevant) between-subjects design. Another 15 participants quit the survey after condition assignment but before completing the dependent variables (four from the first condition, seven from the middle condition, and four from the last condition). We pre-registered our data collection and analysis plan at http://aspredicted.org/blind.php?x=rd9f5j.

Materials and Procedure. Participants in the first, middle, or last condition read the same scenario about a group of scientists as in Studies 4 and 5. However, to vary the group size, in this study we used a group of seven scientists. Participants specifically read that the scientists received their visa to come to the US in an arbitrary order that was determined by their birthdays. The focal member (Milom) received his visa first, fourth (i.e., middle), or last. Participants then read that Milom does a fantastic job and that his work enabled a major breakthrough in research (relevant information condition) or that he is a kind-hearted and generous person (irrelevant information condition).

Next, to assess the effect of the relevant versus irrelevant information on the focal scientist on expectations about the other scientists, we asked participants in the relevant information condition how they thought the other researchers—Jaron, Clovas, Bridi, Erlend, Jago, and Cashel—did on their projects (1 = *very poorly*; 9 = *very well*), α = .94. Participants in

the irrelevant information condition reported how likely they thought the other researchers were kind-hearted and generous (1 = *very unlikely*; 9 = *very likely*), α = .95. Finally, participants responded to the same attention check as in Study 4. Eleven percent of participants (19 from the first condition, and 23 from the middle and last conditions) falsely reported that the best researcher got their visa first. Five percent of participants (13 from the first condition, six from the middle condition, and 11 from the last condition) falsely reported that the researcher whose case was easiest to process got their visa first. These participants thus failed the attention check. Another eight percent of participants (15 from the first condition, 23 from the middle condition, and 11 from the last condition) did not remember how the order had been assigned. We retained both the participants who failed the attention check and those who did not remember the order assignment in the subsequent analyses and note that excluding them did not change the pattern or the significance of the results.

Results and Discussion

An ANOVA of the focal member's position in the sequence × information type on expected group-performance yielded a main effect of position, $F(2, 599) = 6.200$, $p = .002$, $\eta^2 = .02$. Participants perceived the rest of the group as more positive on the respective quality (performance vs. generosity) when the focal member was first, compared to middle or last. Furthermore, we found a significant main effect of information type, $F(1, 599) = 4.583$, $p = .033$, $\eta^2 = .01$, such that the rest of the group was seen as more positive in the relevant information (vs. irrelevant information) condition. Importantly, in support of our hypothesis, we found the predicted interaction, $F(2, 599) = 3.258$, $p = .039$, $\eta^2 = .01$.

In the relevant information condition, we found the first-member heuristic. As expected, when the good performer was in the first position, participants expected the rest of the group to

have performed better (M = 6.58, SD = 1.49) than when the good performer was middle (M = 5.75, SD = 1.19), $p < .001$, CI[0.450, 1.215]. Unexpectedly, the difference between the first and the last (M = 6.26, SD = 1.45) was directional but did not reach significance, $p = .101$, CI[-0.063, 0.700]. However, when the last performer did well, the rest of the group was seen as better than when the middle performer did well, $p = .009$, CI[0.130, 0.898]. Because the middle and last condition did not differ from each other in previous studies, these results are likely due to chance.

In the irrelevant information condition, we did not find the first-member heuristic. As expected, the group was seen as similarly generous when the first member was described as such (M = 6.01, SD = 1.37), compared to the middle (M = 5.88, SD = 1.52), $p = .511$, CI[-0.515, 0.257], or last (M = 5.98, SD = 1.26), $p = .906$, CI[-0.359, 0.405]. There was also no difference between the middle and last condition, $p = .588$, CI[-0.491, 0.279].

This study showed type of information as a boundary condition of the first-member heuristic. People use the first member (vs. middle/last) to make inferences about other members to larger extents when the information is relevant for membership in the group, that is, when it is plausible that the information about the first is actually informative about other members. If the information about the first is likely irrelevant for membership in the group, the first member is no more diagnostic than other members.