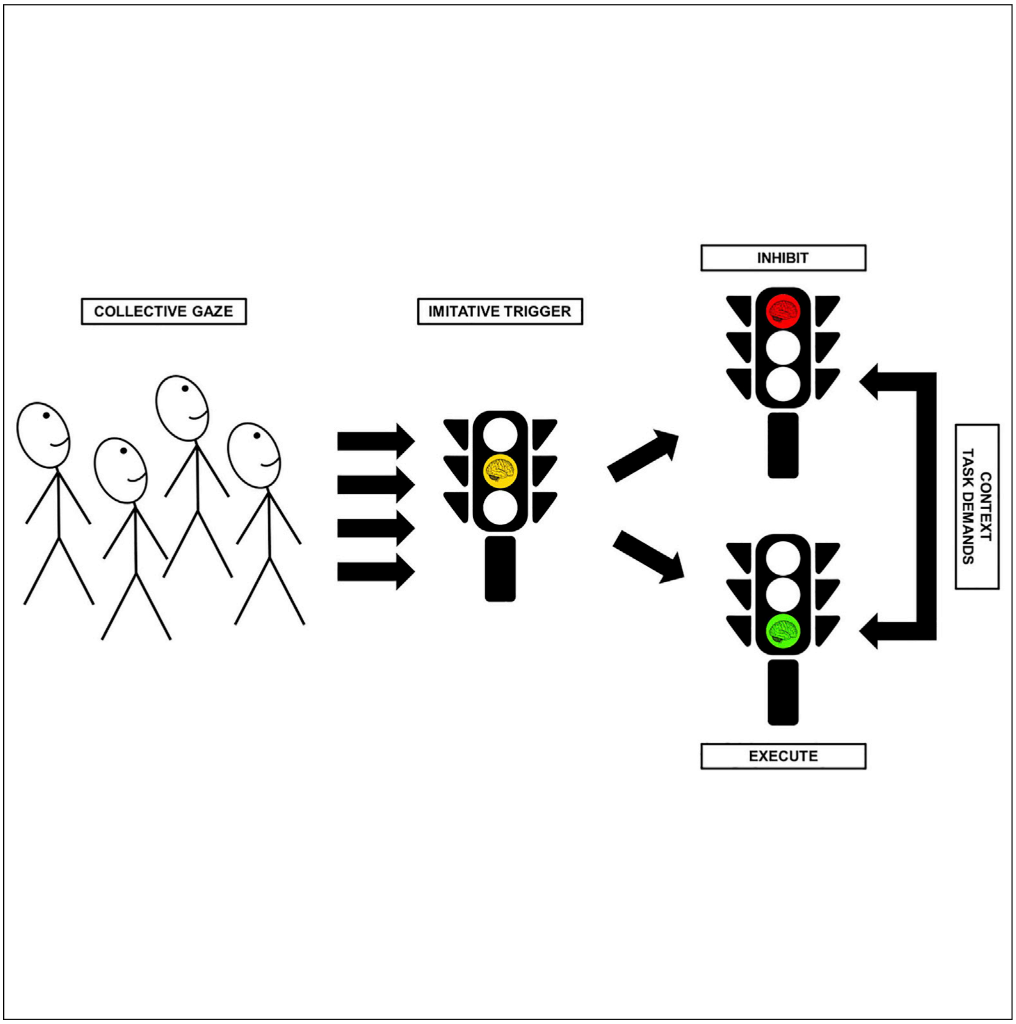
iScience

CellPress
OPEN ACCESS

Article

Evidence for a two-step model of social group influence



Emiel Cracco, Ulysses Bernardet, Robbe Sevenhant, ..., Wouter Durnez, Klaas Bombeke, Marcel Brass

emiel.cracco@ugent.be

Highlights

Groups influence human behavior in two stages

First, groups elicit a reflexive tendency to imitate

Second, strategic processes decide to execute or inhibit this reflex

Social group conformity may be more automatic than previously thought

Cracco et al., iScience 25, 104891
September 16, 2022 © 2022
https://doi.org/10.1016/j.isci.2022.104891

iScience

CellPress
OPEN ACCESS

Article

Evidence for a two-step model of social group influence

Emiel Cracco,[1,2,6,*] Ulysses Bernardet,[3] Robbe Sevenhant,[1] Nette Vandenhouwe,[1] Fran Copman,[1] Wouter Durnez,[4] Klaas Bombeke,[4] and Marcel Brass[1,5]

SUMMARY

Social group influence plays an important role in societally relevant phenomena such as rioting and mass panic. One way through which groups influence individuals is by directing their gaze. Evidence that gaze following increases with group size has typically been explained in terms of strategic processes. Here, we tested the role of reflexive processes. In an ecologically valid virtual reality task, we found that participants were more likely to follow the group's gaze when more people looked, even though they knew the group provided no relevant information. Interestingly, participants also sometimes changed their mind after starting to follow the gaze of the group, indicating that automatic imitation can be overruled by strategic processes. This suggests that social group influence is best explained by a two-step model in which bottom-up imitative processes first elicit a reflexive tendency to imitate, before top-down strategic processes determine whether to execute or inhibit this reflex.

INTRODUCTION

Social group influence is an important driver of human behavior (Latane, 1981). Group dynamics influence whether people break the law (Krause et al., 2021), resort to violence (Hylander and Granström, 2010; Nassauer, 2019), or help others in need (Darley and Latané, 1968; Fischer et al., 2011). It can also improve decision-making (Krause et al., 2010; Tump et al., 2020). One key mechanism through which groups influence behavior is by directing people's gaze and attention (Gallup et al., 2012; Krause et al., 2021). In one of the first studies to make this point, Milgram et al. (1969) asked a group of confederates to walk down a New York City street before suddenly stopping and looking up at a tall building. They then measured how often unaware passersby imitated the group's behavior and found a clear positive relationship between the size of the stimulus group and the probability of looking up (see also Gallup et al., 2012; Jorjafki et al., 2018; Knowles and Bassett, 1976).

The results of Milgram et al. (1969) indicate that collective gaze has a strong pull on attention. What enters our focus of attention, in turn, often determines how we respond. Protesters are, for instance, more likely to resort to violence when they witness offensive behavior from the opposing side (Adang, 2011; Nassauer, 2018). As a result, simply by steering our gaze, groups can already have a strong influence on behavior (Krause et al., 2021). But why do we follow the gaze of groups? According to the traditional account, we do so because it is adaptive: when we see a large group of people looking in the same direction, we assume that they must be looking at something important and therefore decide to follow the group's gaze as a means of obtaining relevant information (Gallup et al., 2012; Jorjafki et al., 2018; Milgram et al., 1969). In other words, the traditional account argues that the influence of groups on attention is driven by top-down, strategic processes (Jorjafki et al., 2018).

At the same time, cognitive research suggests that social group influence may also have a more reflexive, bottom-up component (Cracco et al., 2018a; Frischen et al., 2007). For example, research on gaze cueing has shown that people are faster to detect stimuli that are preceded by a face looking in the direction of the stimulus, even when gaze direction is non-predictive of stimulus location (Driver et al., 1999; Friesen and Kingstone, 1998) or is predictive of a stimulus in the opposite direction (Driver et al., 1999). More generally, research on automatic imitation has shown that people spontaneously imitate a wide range of behaviors (Chartrand and Bargh, 1999; Genschow et al., 2017) even when doing so is disadvantageous (Brass et al., 2000; Stürmer et al., 2000). Both gaze cueing (Capozzi et al., 2015, 2021; Sun et al., 2017) and automatic

[1]Department of Experimental Psychology, Ghent University, Gent, Belgium

[2]Department of Experimental Clinical and Health Psychology, Ghent University, Gent, Belgium

[3]Aston Institute of Urban Technology and the Environment, Aston University, Birmingham, UK

[4]imec-mict-UGent, Department of Communication Sciences, Ghent University, Gent, Belgium

[5]Berlin School of Mind and Brain, Humboldt University of Berlin, Berlin, Germany

[6]Lead contact

*Correspondence: emiel.cracco@ugent.be

https://doi.org/10.1016/j.isci.2022.104891

Figure 1. Example frames of the experiment
The left frame shows the virtual agents looking up. The right frame shows the participant's response. See Video S1 for an example video.

imitation (Cracco et al., 2015; Cracco and Brass, 2018a, 2018b) have recently been shown to increase with group size, presumably because larger groups elicit more motor resonance (Cracco et al., 2016, 2019). This suggests that the influence of groups on attention may also have a more reflexive, bottom-up component.

However, studies on gaze cueing and automatic imitation are conducted in the lab, with artificial computer tasks that only indirectly measure conformity. As a result, the findings of these studies are difficult to translate to social group influence in the real world. Conversely, field studies, like the one by Milgram et al. (1969), are close to real-life situations, but often lack the experimental control needed to isolate specific processes. Hence, whether not only just strategic but also reflexive processes contribute to social group influence is still unknown. Here, we address this question by using virtual reality (VR) to increase ecological validity while retaining experimental control. Specifically, we created an immersive VR task based on the study of Milgram et al. (1969), but following the structure of gaze cueing and automatic imitation experiments. In our task, individuals watched an open-air film together with 10 virtual agents (Figure 1). Every so often, a sound was played, requesting participants to look up to the left (forced choice), to look up to the right (forced choice), or to randomly choose where to look with a 50/50 ratio (free choice). At the same time, a variable number of agents (1–10) also looked up to the left or right. Importantly, to complete the trial, participants had to detect a fire they knew was always present in both locations. As a result, the gaze direction of the virtual agents contained no information about the target location, and we could distill the influence of reflexive, bottom-up processes from indirect as well as direct measures of gaze following.

The indirect measure of gaze following was obtained from forced choice trials and was defined as faster responses on trials where participants had to look in the same direction as the virtual agents (congruent trials) than on trials where they had to look in a different direction (incongruent trials; Cracco et al., 2018a; Frischen et al., 2007). In contrast, the direct measure of gaze following was obtained from free choice trials and was defined as the proportion of trials in which participants decided to follow the agents (Gallup et al., 2012; Jorjafki et al., 2018; Milgram et al., 1969). If not only strategic processes but also reflexive processes contribute to social group influence, gaze following should increase with the number of virtual agents looking up even when the group contains no relevant information. This should be true for both types of gaze following measured here. In addition, we might also see similar effects in other aspects of participants' behavior, such as movement time (i.e., how long it takes to execute the movement), errors (i.e., forced choice trials where participants look in the wrong direction), partial errors (i.e., forced choice trials where participants first look in the wrong direction but then correct themselves), and partial choices (i.e., free choice trials where participants first look in one direction but then correct to the other direction).

RESULTS

Virtual presence ratings

To assess whether participants felt present in the virtual environment, we administered the Igroup Presence Questionnaire (IPQ; Schubert et al., 2001). As can be seen from Table 1, this revealed that ratings of general and spatial presence were relatively high, whereas ratings of involvement and realism were somewhat lower, presumably due to the repetitive nature of the experiment and the animated environment. When added as a predictor to the statistical models, none of the four scales significantly influenced the effects of group size on behavior.

CellPress
OPEN ACCESS

Table 1. Descriptive results of the Igroup Presence Questionnaire (IPQ)

Subscale	M	SD	α
General Presence	5.20	1.17	N/A
Spatial Presence	5.33	0.84	0.64
Involvement	3.64	1.34	0.82
Experienced Realism	3.48	1.15	0.79

Items on the IPQ are scored on a Likert scale ranging from 1 to 7. The General Presence scale includes only one item. As a result, Cronbach's alpha cannot be calculated. See STAR Methods for details.

Forced choice results

The influence of the group on forced choice behavior was measured using (generalized) linear mixed effects models with group size (1–10) as a continuous predictor and congruency (congruent vs. incongruent) as a factor. In addition, based on previous evidence that group size tends to have an asymptotic influence on gaze following (Capozzi et al., 2018; Gallup et al., 2012; Jorjafki et al., 2018; Milgram et al., 1969), we also exploratively looked for non-linear trends in the data by comparing each linear model to an equivalent non-linear model.

Reaction times

The reaction time analysis revealed a main effect of group size, $t(20211) = 7.62$, $p < 0.001$, with faster responses as group size increased, a main effect of congruency, $t(149) = 13.19$, $p < 0.001$, with faster responses on congruent than on incongruent trials, and a group size × congruency interaction, $t(20221) = 4.41$, $p < 0.001$, with a larger congruency effect as group size increased. Comparing the linear to the non-linear model further revealed a significantly better fit for the non-linear model, $\chi^2(2) = 6.63$, $p = 0.036$, indicating that the congruency effect in reaction times changed with group size according to a slightly asymptotic curve (Figure 2).

Movement times

The movement time analysis revealed a main effect of group size, $t(20340) = 2.85$, $p = 0.004$, with slower movements as group size increased, and a main effect of congruency, $t(20340) = 3.02$, $p = 0.003$, with faster movements on congruent than on incongruent trials, but no group size × congruency interaction, $t(20340) = 0.36$, $p = 0.720$. Comparing the linear model to the non-linear model revealed no significant difference, $\chi^2(2) = 3.79$, $p = 0.150$ (Figure 2).

Partial errors

The partial error analysis revealed no main effect of group size, $z = 0.58$, $p = 0.564$, but did reveal a main effect of congruency, $z = 8.07$, $p < 0.001$, with fewer partial errors on congruent than on incongruent trials, and a group size × congruency interaction, $z = 3.10$, $p = 0.002$, showing that this congruency effect increased with group size. Comparing the linear to the non-linear model revealed a significantly better fit for the non-linear model, $\chi^2(2) = 10.68$, $p = 0.005$, suggesting an asymptotic curve for the effect of group size on the congruency effect (Figure 2).

Error rates

The error rate analysis revealed a main effect of group size, $z = 3.68$, $p < 0.001$, with more errors as group size increased, a main effect of congruency, $z = 10.79$, $p < 0.001$, with fewer errors on congruent than on incongruent trials, and a group size × congruency interaction, $z = 3.51$, $p < 0.001$, with a larger congruency effect as group size increased. Comparing the linear to the non-linear model revealed no significant difference, $\chi^2(2) = 0.69$, $p = 0.709$ (Figure 2).

Free choice results

The influence of group size on free choice reaction times, movement times, and partial choices was measured using (generalized) linear mixed effects models with group size (1–10) as a continuous predictor and choice (follow vs. not follow) as a factor. The influence on follow choices was measured with a generalized linear mixed effects model including only group size as predictor. In line with the forced choice analyses, these models were again compared to non-linear models to test whether the effect of group size followed an asymptotic curve.

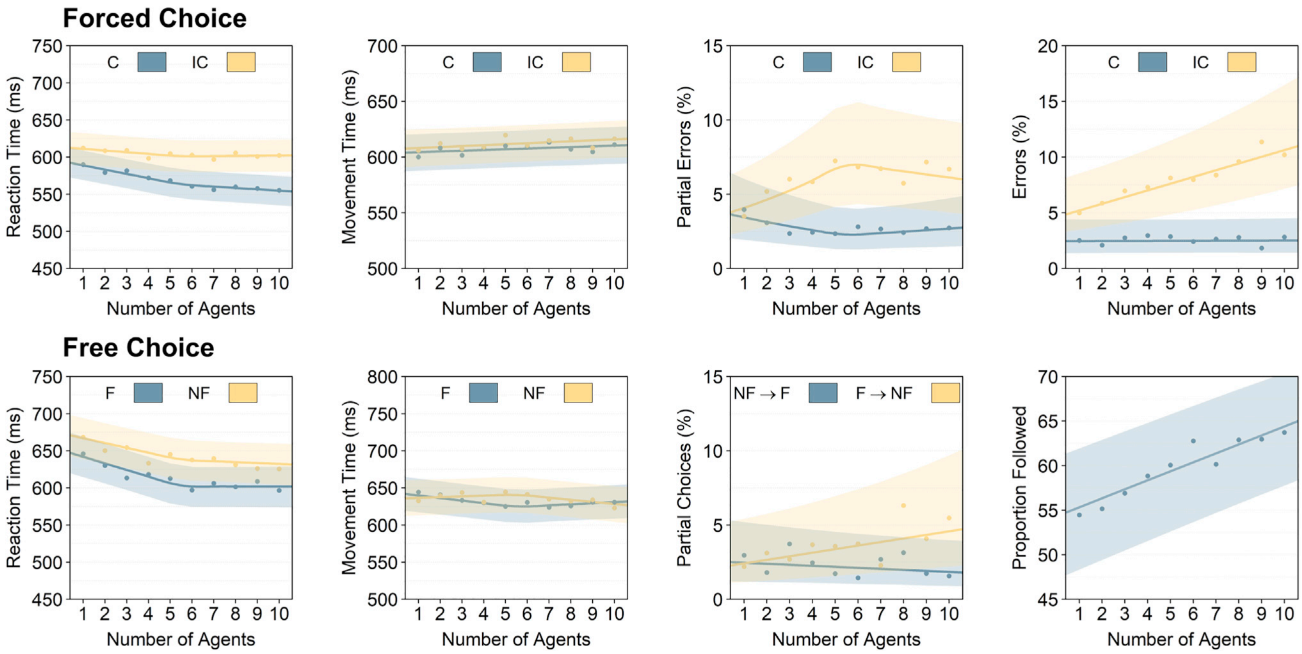


Figure 2. Forced and free choice results

The data were fitted using both linear and non-linear models. The shown fit lines and error bands reflect the best fitting model. Error bands are 80% prediction intervals fitted using the merTools package in R (Knowles and Frederick, 2020), showing the interval in which 80% of new observations (default package value) are expected to fall according to the model. Note that the % of partial choices is calculated with respect to the eventually chosen target. C: congruent, IC: incongruent, F: follow, NF: Not Follow.

Reaction times

The reaction time analysis revealed a main effect of group size, $t(10445) = 7.81$, $p < 0.001$, with faster choices as group size increased, and a main effect of choice, $t(10467) = 10.17$, $p < 0.001$, with faster choices when participants decided to follow the agents, but no group size × choice interaction, $t(10449) = 0.46$, $p = 0.644$. Comparing the linear to the non-linear model revealed a significantly better fit for the non-linear model, $\chi^2(2) = 15.10$, $p < 0.001$, indicating that the effect of group size on reaction time followed an asymptotic curve (Figure 2).

Movement times

The movement time analysis revealed a main effect of group size, $t(10440) = 2.24$, $p = 0.025$, with faster movements as group size increased, and a main effect of choice, $t(130) = 2.24$, $p = 0.027$, with faster movements when participants followed the group, but no group size × choice interaction, $t(10370) = 0.13$, $p = 0.897$. Comparing the linear to the non-linear model revealed a significantly better fit for the non-linear model, $\chi^2(2) = 7.57$, $p = 0.023$, suggesting an unexpected reverse U-shaped pattern for the effect of group size on the congruency effect (Figure 2).

Partial choices

The partial choice analysis revealed no main effects of group size, $z = 0.98$, $p = 0.328$, or choice, $z = 1.82$, $p = 0.069$. However, the main effect of choice was close to significance, indicating that participants tended to more often correct their response when they initially followed the group. This tendency was further qualified by a significant group size × choice interaction, $z = 2.61$, $p = 0.009$, indicating that it became stronger as group size increased. Comparing the linear and non-linear model revealed no significant difference, $\chi^2(2) = 0.25$, $p = 0.882$ (Figure 2).

Follow choices

The follow choice analysis revealed that participants followed the agents more than would be expected based on chance, $z = 9.25$, $p < 0.001$, and additionally revealed a main effect of group size, $z = 6.15$, $p < 0.001$, indicating that the probability of following the agents increased with group size. Comparing the linear and non-linear model revealed no significant difference, $\chi^2(2) = 2.88$, $p = 0.090$, although there was a non-significant tendency toward an asymptotic curve (Figure 2).

DISCUSSION

What we attend to strongly determines what we do. Accordingly, by directing our attention and gaze (Gallup et al., 2012; Jorjafki et al., 2018; Milgram et al., 1969), groups can have a strong influence on behavior (Latane, 1981), contributing to acts of violence (Adang, 2011; Krause et al., 2021; Nassauer, 2018), helping behavior (Darley and Latané, 1968; Fischer et al., 2011), and other societally relevant deeds. Yet what drives the influence of groups on gaze following is not yet clear. According to the top-down account, large groups direct our gaze because we reason that when many people look in the same direction, they must be looking at something important (Gallup et al., 2012; Jorjafki et al., 2018; Milgram et al., 1969). In contrast, the bottom-up account argues that larger groups simply provide a stronger trigger to the motor system and therefore elicit a stronger urge to imitate (Cracco et al., 2015, 2016; Cracco and Brass, 2018b). Here, we tested the bottom-up account by designing a VR task that probed gaze following of groups that contained no relevant information about reality.

Both direct and indirect measures of gaze following increased with group size. In line with the bottom-up account, this suggests that seeing multiple people perform the same action leads to an automatic imitative response tendency (Cracco et al., 2015; Cracco and Brass, 2018b) that makes it more likely that observers will follow the observed behavior. Crucially, however, such a bottom-up explanation does not necessarily exclude top-down processes in social group influence. Indeed, previous research has provided clear evidence for top-down processes. For example, Jorjafki et al. (2018) found that participants sometimes did not imitate even though they indicated afterward to have noticed people looking up. Similarly, Gallup et al. (2012) found that people tended to inhibit gaze following when the group could see them, causing them to imitate more when they were behind or to the side of the group than when they were in front of it. These data thus suggest that social group influence is best explained by a two-step model, in which reflexive imitative processes first elicit an initial tendency to imitate, before more deliberate processes decide, based on an interpretation of the social situation, whether to execute or inhibit the imitative trigger (Cracco et al., 2015, 2016). Such a process model is consistent with brain imaging evidence on imitation, which has shown that the decision to overtly imitate relies on a gating mechanism that regulates automatic motor resonance elicited by observing other people's actions (Bien et al., 2009).

A two-step model is also supported by the data of the current study. That is, both on forced and on free choice trials, we found clear evidence for partial responses, where participants first moved their head in one direction but then corrected to the other direction. On free choice trials, a partial response reflects a change of intention (Furstenberg et al., 2015), suggesting that the initial decision to follow a group can be overruled by slower top-down processes. In the current study, these top-down influences were likely driven by strategic processes induced by the task instructions. More specifically, the instruction to look roughly equally often in both directions likely caused participants to sometimes overturn a reflexive imitative response when imitation would cause them to look in a direction they had already looked in often or recently. Indeed, research has shown that humans often misconceive randomness, in the sense that they see repetitions as inconsistent with random behavior (Bar-Hillel and Wagenaar, 1991). In real life, changes of intention are more likely driven by interpretative processes (Cracco et al., 2015, 2016). That is, studies suggest that people tend to inhibit imitation of inappropriate behavior (Cracco et al., 2018b) and are influenced more by leaders (Capozzi et al., 2016) and previously reliable individuals (Capozzi et al., 2021), even to the extent that the influence of the majority group is sometimes overruled (Capozzi et al., 2021). More generally, a two-step model is consistent with recent evidence that gaze following is the manifestation of a complex interplay between basic attentional and advanced social processes (e.g., Capozzi and Ristic, 2021; Colombatto et al., 2020; Guterstam et al., 2019; Mayrand et al., 2021).

Moving beyond imitative responses, we also found evidence for a more general influence of groups on behavior. Specifically, we found that increases in group size led to a decrease in reaction time, likely reflecting a social facilitation effect (Zajonc, 1965), whereby a large group of people making a sudden movement has an invigorating effect that primes the motor system for action. Interestingly, on forced choice trials, this decrease in reaction time as a function of group size was offset by a corresponding increase in movement time. This is consistent with the proposed two-step model, distinguishing between initial bottom-up processes, captured mainly by reaction times, and slower top-down processes, captured mainly by movement times. In particular, it suggests that motor inhibition is applied following fast responses to prevent premature and potentially erroneous decisions (Cracco and Brass, 2018b; Cracco and Cooper, 2019). Although speculative, such a mechanism might also explain why the same dissociation between reaction and movement times was not found on free choice trials. On free choice trials, no errors could be made, and hence there was no longer any need to inhibit fast responses. Importantly, however, this is a post-hoc explanation that cannot be directly supported by the data and will therefore have to be confirmed in future research.

Finally, in exploratory analyses, we attempted to replicate evidence that the relationship between group size and gaze following is asymptotic (Capozzi et al., 2018; Gallup et al., 2012; Jorjafki et al., 2018; Milgram et al., 1969). Supporting this idea, group size had an asymptotic influence on reaction times and partial errors. However, it had a more linear influence on the other measures. Similarly, previous research also found asymptotic relationships for some but not all types of behavior. For example, Milgram et al. (1969) found that group size had an asymptotic influence on people looking up, but a linear influence on people stopping to look up. Gallup et al. (2012) further found that an asymptote might only be reached for larger groups with more than 10 members. This suggests that the curve of the relationship between group size and gaze following depends on the type of behavior being measured and on the social context. However, which factors exactly determine the curve is not yet clear.

To conclude, this study shows that the influence of groups on gaze following relies at least partly on reflexive, bottom-up processes. Specifically, it suggests that the drawing power of groups (Gallup et al., 2012; Jorjafki et al., 2018; Milgram et al., 1969) is best explained by a two-stage model in which reflexive processes first elicit an initial tendency to imitate, before strategic processes determine whether to act on this tendency. By showing how cognitive processes at the individual level determine behavior at the group level, these findings have important implications for understanding collective behavior in biology (e.g., Couzin, 2018; Sumpter, 2006), psychology (e.g., Raafat et al., 2009; Tump et al., 2020), and sociology (e.g., Hylander and Granström, 2010; Nassauer, 2019), and may advance our understanding of how societally relevant phenomena such as rioting or mass panic unfold (Krause et al., 2021).

Limitations of the study

Like all studies, the current study also has a number of limitations. First, the rating data showed that even though participants felt present in the virtual environment, they did not feel much immersed. Given that the degree to which the virtual environment resembles a real environment has previously been shown to influence social behavior (Durnez et al., 2020), an important task for future work will be to further improve the virtual scenarios. Second, the current study investigated only one type of social behavior: gaze following. By determining what we see, gaze following plays an important role in shaping what we do. Nevertheless, an important remaining question is whether a similar two-step model can also be applied to other types of social behavior such as conformity (Toelch and Dolan, 2015) or helping (Fischer et al., 2011).

STAR★METHODS

Detailed methods are provided in the online version of this paper and include the following:

- KEY RESOURCES TABLE
- RESOURCE AVAILABILITY
  - Lead contact
  - Materials availability
  - Data and code availability
- EXPERIMENTAL MODEL AND SUBJECT DETAILS
- METHOD DETAILS
  - Materials
  - Task and procedure
- QUANTIFICATION AND STATISTICAL ANALYSIS

SUPPLEMENTAL INFORMATION

Supplemental information can be found online at https://doi.org/10.1016/j.isci.2022.104891.

ACKNOWLEDGMENTS

E.C. was supported by a Research Foundation Flanders postdoctoral fellowship (FWO18/PDO/049 and 12U0322N). M.B. was supported by an Einstein Strategic Professorship of the Einstein Foundation Berlin and by the Deutsche Forschungsgemeinschaft (DFG, German Research Foundation) under Germany's Excellence Strategy – EXC 2002/1 "Science of Intelligence" – project number 390523135.

AUTHOR CONTRIBUTIONS

Conceptualization: E.C. and M.B.; Methodology: E.C., U.B., and M.B.; Software: E.C., U.B., and R.S.; Validation: E.C., F.C., W.D., and K.B.; Formal Analysis: E.C.; Investigation: F.C., R.S., and N.V.; Writing – Original Draft: E.C.; Writing – Review & Editing: U.B., W.D., K.B., and M.B.; Visualization: E.C.

DECLARATION OF INTERESTS

The authors declare no competing interests.

Received: February 21, 2022
Revised: May 11, 2022
Accepted: August 3, 2022
Published: September 16, 2022

REFERENCES

Adang, O.M.J. (2011). Initiation and escalation of collective violence: an observational study. In Preventing crowd violence (Lynne Rienner Publishers), pp. 47–68.

Arrington, C.M., and Logan, G.D. (2004). The cost of a voluntary task switch. Psychol. Sci. *15*, 610–615. https://doi.org/10.1111/j.0956-7976.2004.00728.x.

Baayen, R.H., Davidson, D.J., and Bates, D.M. (2008). Mixed-effects modeling with crossed random effects for subjects and items. J. Mem. Lang. *59*, 390–412. https://doi.org/10.1016/j.jml.2007.12.005.

Bar-Hillel, M., and Wagenaar, W.A. (1991). The perception of randomness. Adv. Appl. Math. *12*, 428–454. https://doi.org/10.1016/0196-8858(91)90029-I.

Bates, D., Mächler, M., Bolker, B., and Walker, S. (2014). Fitting linear mixed-effects models using lme4. J. Stat. Software *67*.

Bien, N., Roebroeck, A., Goebel, R., and Sack, A.T. (2009). The brain's intention to imitate: the neurobiology of intentional versus automatic imitation. Cerebr. Cortex *19*, 2338–2351. https://doi.org/10.1093/cercor/bhn251.

Brass, M., Bekkering, H., Wohlschläger, A., Prinz, W., Wohlschläger, A., and Prinz, W. (2000). Compatibility between observed and executed finger movements: comparing symbolic, spatial, and imitative cues. Brain Cognit. *44*, 124–143. https://doi.org/10.1006/brcg.2000.1225.

Capozzi, F., Bayliss, A.P., Elena, M.R., and Becchio, C. (2015). One is not enough: group size modulates social gaze-induced object desirability effects. Psychon. Bull. Rev. *22*, 850–855. https://doi.org/10.3758/s13423-014-0717-z.

Capozzi, F., Bayliss, A.P., and Ristic, J. (2018). Gaze following in multiagent contexts: evidence for a quorum-like principle. Psychon. Bull. Rev. *25*, 2260–2266. https://doi.org/10.3758/s13423-018-1464-3.

Capozzi, F., Bayliss, A.P., and Ristic, J. (2021). Standing out from the crowd: both cue numerosity and social information affect attention in multi-agent contexts. Q. J. Exp. Psychol. *74*, 1737–1746. https://doi.org/10.1177/17470218211013028.

Capozzi, F., Becchio, C., Willemse, C., and Bayliss, A.P. (2016). Followers are not followed: observed group interactions modulate subsequent social attention. J. Exp. Psychol. Gen. *145*, 531–535. https://doi.org/10.1037/xge0000167.

Capozzi, F., and Ristic, J. (2021). Attentional gaze dynamics in group interactions. Vis. Cognit. *30*, 135–150. https://doi.org/10.1080/13506285.2021.1925799.

Chartrand, T.L., and Bargh, J.A. (1999). The chameleon effect: the perception-behavior link and social interaction. J. Pers. Soc. Psychol. *76*, 893–910.

Colombatto, C., Chen, Y.-C., and Scholl, B.J. (2020). Gaze deflection reveals how gaze cueing is tuned to extract the mind behind the eyes. Proc. Natl. Acad. Sci. USA *117*, 19825–19829. https://doi.org/10.1073/pnas.2010841117.

Couzin, I.D. (2018). Synchronization: the key to effective communication in animal collectives. Trends Cognit. Sci. *22*, 844–846. https://doi.org/10.1016/j.tics.2018.08.001.

Cracco, E., Bardi, L., Desmet, C., Genschow, O., Rigoni, D., De Coster, L., Radkova, I., Deschrijver, E., and Brass, M. (2018a). Automatic imitation: a meta-analysis. Psychol. Bull. *144*, 453–500. https://doi.org/10.1037/bul0000143.

Cracco, E., and Brass, M. (2018a). Automatic imitation of multiple agents: simultaneous or random representation? J. Exp. Psychol. Hum. Percept. Perform. *44* (5), 729–740. https://doi.org/10.1037/xhp0000489.

Cracco, E., and Brass, M. (2018b). The role of sensorimotor processes in social group contagion. Cognit. Psychol. *103*, 23–41.

Cracco, E., and Cooper, R.P. (2019). Automatic imitation of multiple agents: a computational model. Cognit. Psychol. *113*, 101224. https://doi.org/10.1016/j.cogpsych.2019.101224.

Cracco, E., De Coster, L., Andres, M., and Brass, M. (2015). Motor simulation beyond the dyad: automatic imitation of multiple actors. J. Exp. Psychol. Hum. Percept. Perform. *41*, 1488–1501. https://doi.org/10.1037/a0039737.

Cracco, E., De Coster, L., Andres, M., and Brass, M. (2016). Mirroring multiple agents: motor resonance during action observation is modulated by the number of agents. Soc. Cognit. Affect Neurosci. *11*, 1422–1427. https://doi.org/10.1093/scan/nsw059.

Cracco, E., Genschow, O., Radkova, I., and Brass, M. (2018b). Automatic imitation of pro- and antisocial gestures: is implicit social behavior censored? Cognition *170*, 179–189. https://doi.org/10.1016/j.cognition.2017.09.019.

Cracco, E., Keysers, C., Clauwaert, A., and Brass, M. (2019). Representing multiple observed actions in the motor system. Cerebr. Cortex *29*, 3631–3641. https://doi.org/10.1093/cercor/bhy237.

Latané, B., and Darley, J.M. (1968). Group inhibition of bystander intervention in emergencies. J. Pers. Soc. Psychol. *10*, 215–221. https://doi.org/10.1037/h0026570.

Driver, J., Davis, G., Ricciardelli, P., Kidd, P., Maxwell, E., and Baron-Cohen, S. (1999). Gaze perception triggers reflexive visuospatial orienting. Vis. Cognit. *6*, 509–540. https://doi.org/10.1080/135062899394920.

Durnez, W., Bombeke, K., Joundi, J., Zheleva, A., Cracco, E., Copman, F., Brass, M., Saldien, J., and De Marez, L. (2020). Fake people, real effects. In Virtual, Augmented and Mixed Reality. Design and Interaction, J.Y.C. Chen and G. Fragomeni, eds. (Springer International Publishing), pp. 440–452. https://doi.org/10.1007/978-3-030-49695-1_29.

Fischer, P., Krueger, J.I., Greitemeyer, T., Vogrincic, C., Kastenmüller, A., Frey, D., Heene, M., Wicher, M., and Kainbacher, M. (2011). The bystander-effect: a meta-analytic review on bystander intervention in dangerous and non-dangerous emergencies. Psychol. Bull. *137*, 517–537. https://doi.org/10.1037/a0023304.

Friesen, C.K., and Kingstone, A. (1998). The eyes have it! Reflexive orienting is triggered by nonpredictive gaze. Psychon. Bull. Rev. *5*, 490–495. https://doi.org/10.3758/BF03208827.

Frischen, A., Bayliss, A.P., and Tipper, S.P. (2007). Gaze cueing of attention: visual attention, social cognition, and individual differences. Psychol. Bull. *133*, 694–724. https://doi.org/10.1037/0033-2909.133.4.694.

Furstenberg, A., Breska, A., Sompolinsky, H., and Deouell, L.Y. (2015). Evidence of change of intention in picking situations. J. Cognit. Neurosci. *27*, 2133–2146. https://doi.org/10.1162/jocn_a_00842.

Gallup, A.C., Hale, J.J., Sumpter, D.J.T., Garnier, S., Kacelnik, A., Krebs, J.R., and Couzin, I.D. (2012). Visual attention and the acquisition of information in human crowds. Proc. Natl. Acad. Sci. USA *109*, 7245–7250. https://doi.org/10.1073/pnas.1116141109.

Genschow, O., Van Den Bossche, S., Cracco, E., Bardi, L., Rigoni, D., Brass, M., Van Den Bossche, S., Cracco, E., Bardi, L., Rigoni, D., and Brass, M. (2017). Mimicry and automatic imitation are not correlated. PLoS One *12*. e0183784–21. https://doi.org/10.1371/journal.pone.0183784.

Guterstam, A., Kean, H.H., Webb, T.W., Kean, F.S., and Graziano, M.S.A. (2019). Implicit model of other people's visual attention as an invisible, force-carrying beam projecting from the eyes. Proc. Natl. Acad. Sci. USA *116*, 328–333. https://doi.org/10.1073/pnas.1816581115.

Halas, J., and Batchelor, J. (1954). Animal Farm (Hallas and Batchelor).

Hastie, T., Tibshirani, R., and Friedman, J. (2009). The Elements of Statistical Learning, 2nd ed. (New York: Springer-Verlag).

Hylander, I., and Granström, K. (2010). Organizing for a peaceful crowd: an example of a Football match. Forum Qualitative Sozialforschung/Forum Qual. Soc. Res. *11*. https://doi.org/10.17169/fqs-11.2.1462.

Jorjafki, E.M., Sagarin, B.J., and Butail, S. (2018). Drawing power of virtual crowds. J. R. Soc. Interface *15*, 20180335. https://doi.org/10.1098/rsif.2018.0335.

Knowles, E.S., and Bassett, R.L. (1976). Groups and crowds as social entities: effects of activity, size, and member similarity on nonmembers. J. Pers. Soc. Psychol. *34*, 837–845. https://doi.org/10.1037/0022-3514.34.5.837.

Knowles, J.E., and Frederick, C. (2020). merTools: tools for analyzing mixed effect regression models (0.5.2) [R]. https://CRAN.R-project.org/package=merTools.

Krause, J., Romanczuk, P., Cracco, E., Arlidge, W., Nassauer, A., and Brass, M. (2021). Collective rule breaking. Trends Cognit. Sci.

Krause, J., Ruxton, G.D., and Krause, S. (2010). Swarm intelligence in animals and humans. Trends Ecol. Evol. *25*, 28–34. https://doi.org/10.1016/j.tree.2009.06.016.

Kumle, L., Võ, M.L.-H., and Draschkow, D. (2021). Estimating power in (generalized) linear mixed models: an open introduction and tutorial in R. Behav. Res. Methods *53*, 2528–2543. https://doi.org/10.3758/s13428-021-01546-0.

Kuznetsova, A., Brockhoff, P.B., and Christensen, R.H.B. (2017). lmerTest package: tests in linear mixed effects models. J. Stat. Software *82*, 1–26. https://doi.org/10.18637/jss.v082.i13.

Latane, B. (1981). The psychology of social impact. Am. Psychol. *36*, 343–356. https://doi.org/10.1037/0003-066X.36.4.343.

Matuschek, H., Kliegl, R., Vasishth, S., Baayen, H., and Bates, D. (2017). Balancing Type I error and power in linear mixed models. J. Mem. Lang. *94*, 305–315. https://doi.org/10.1016/j.jml.2017.01.001.

Mayrand, F., McCrackin, S., Capozzi, F., and Ristic, J. (2021). Gaze communicates both cue direction and agent mental states. Preprint at PsyArXiv. https://doi.org/10.31234/osf.io/9asy8.

Milgram, S., Bickman, L., and Berkowitz, L. (1969). Note on the drawing power of crowds of different size. J. Pers. Soc. Psychol. *13*, 79–82. https://doi.org/10.1037/h0028070.

Nassauer, A. (2018). Situational dynamics and the emergence of violence in protests. Psychol. Violence *8*, 293–304. https://doi.org/10.1037/vio0000176.

Nassauer, A. (2019). Situational Breakdowns: Understanding Protest Violence and Other Surprising Outcomes (Oxford University Press). https://doi.org/10.1093/oso/9780190922061.001.0001.

Pan, X., and Hamilton, A.F.d.C. (2018). Why and how to use virtual reality to study human social interaction: the challenges of exploring a new research landscape. Br. J. Psychol. *109*, 395–417. https://doi.org/10.1111/bjop.12290.

Raafat, R.M., Chater, N., and Frith, C. (2009). Herding in humans. Trends Cognit. Sci. *13*, 420–428. https://doi.org/10.1016/j.tics.2009.08.002.

Schubert, T., Friedmann, F., and Regenbrecht, H. (2001). The experience of presence: factor analytic insights. Presence *10*, 266–281. https://doi.org/10.1162/105474601300343603.

Stürmer, B., Aschersleben, G., Prinz, W., Sturmer, B., Aschersleben, G., and Prinz, W. (2000). Correspondence effects with manual gestures and postures: a study of imitation. J. Exp. Psychol. Hum. Percept. Perform. *26* (6), 1746–1759. https://doi.org/10.1037//0096-1523.26.6.1746.

Sumpter, D.J.T. (2006). The principles of collective animal behaviour. Philos. Trans. R. Soc. Lond. B Biol. Sci. *361*, 5–22. https://doi.org/10.1098/rstb.2005.1733.

Sun, Z., Yu, W., Zhou, J., and Shen, M. (2017). Perceiving crowd attention: gaze following in human crowds with conflicting cues. Atten. Percept. Psychophys. *79* (4), 1039–1049. https://doi.org/10.3758/s13414-017-1303-z.

Toelch, U., and Dolan, R.J. (2015). Informational and normative influences in conformity from a neurocomputational perspective. Trends Cognit. Sci. *19*, 579–589. https://doi.org/10.1016/j.tics.2015.07.007.

Tump, A.N., Pleskac, T.J., and Kurvers, R.H.J.M. (2020). Wise or mad crowds? The cognitive mechanisms underlying information cascades. Sci. Adv. *6*, eabb0266. https://doi.org/10.1126/sciadv.abb0266.

Zajonc, R.B. (1965). Social facilitation. Science *149*, 269–274.

STAR★METHODS

KEY RESOURCES TABLE

REAGENT or RESOURCE	SOURCE	IDENTIFIER
Deposited data		
Behavioral data	This paper	Open Science Framework: https://doi.org/10.17605/OSF.IO/MNEUY
Behavioral data R code	This paper	Open Science Framework: https://doi.org/10.17605/OSF.IO/MNEUY
Software and algorithms		
Unity v2019.4.11f1	Unity Technologies	https://unity.com/
R v4.1.2	R Core Team	https://www.r-project.org/

RESOURCE AVAILABILITY

Lead contact

Further information and requests should be directed to and will be fulfilled by the lead contact, Emiel Cracco (emiel.cracco@ugent.be).

Materials availability

This study did not generate new unique reagents.

Data and code availability

- Behavioral data have been deposited at the Open Science Framework and are publicly available as of the date of publication. DOIs are listed in the key resources table.
- All original code has been deposited at the Open Science Framework and is publicly available as of the date of publication. DOIs are listed in the key resources table.
- Any additional information required to reanalyze the data reported in this paper is available from the lead contact upon request.

EXPERIMENTAL MODEL AND SUBJECT DETAILS

Two samples were collected with the same preregistration (https://aspredicted.org/blind.php?x=cu9ma6). Our initial goal was to collect a single sample with N = 100. This sample size was based on time and resource constraints, but a sensitivity analysis using the mixedpower package in R (Kumle et al., 2021) indicated that it provided us with good power (>90%) to obtain similar effects for the primary dependent variables as those obtained in a pilot study (see Data S1). Unfortunately, during data collection, we discovered a minor error in the experimental program (The SOA variable was mistakenly coded as *soa_means_ms*, whereas the program expected *soa_mean_ms* and consequently ignored the SOA manipulation, instead using an SOA of 0 throughout the entire experiment), as a result of which the delay between the stimulus sound and the agents' movements (stimulus onset asynchrony; SOA) was not manipulated as preregistered. We therefore decided to terminate data collection prematurely and start over. Sample 1 consisted of 60 participants (47 female, 10 male, 3 unknown gender, M_{age} = 18.67, SD_{age} = 1.53, $range_{age}$ = 18–27). Sample 2 consisted of 100 participants (76 female, 24 male, M_{age} = 19.17, SD_{age} = 2.03, $range_{age}$ = 18–31). We had no predictions about SOA, as it was included only to make the responses of the virtual agents slightly more realistic. Hence, we decided to take advantage of our increased sample size by analyzing both samples together. Importantly, however, analyzing the data of the two samples separately revealed very similar results in both samples with little difference between them (Data S2).

All participants were students with normal or corrected-to-normal vision who participated in return for course credit. Participants signed an informed consent before the start of the experiment and were informed that a common side effect of VR is that some people get dizzy or nauseous (Pan and Hamilton, 2018). This happened for 4 participants. For these participants, the experiment was terminated prematurely and only the blocks before the block in which the experiment was stopped were included in the analysis.

The study was approved by the Institutional Ethics Committee of the Faculty of Psychology and Educational Sciences at Ghent University (2020/124).

METHOD DETAILS

Materials

The experiment was programmed in Unity version 2019.4.11f1 and was performed on a desktop PC with an Intel Core i9-9900K CPU, 64GB RAM, and a GeForce RTX 2080 Ti graphics card. An HTC Vive Pro head mounted display (HMD) with built-in headphones was used, offering a 110° field of view with a resolution of 1440 × 1600 px per eye at a refresh rate of 90 Hz.

Task and procedure

After entering the test room, participants first signed an informed consent and were given information about the experiment. Next, they put on the HMD and completed a practice phase of 9 trials in which accuracy feedback was provided in the form of a 'ping' (correct) or 'buzz' (error) sound. On free choice trials, all responses were considered correct. The practice phase was followed by the experiment proper, which consisted of 4 blocks of 66 trials without accuracy feedback. Participants performed the task standing up but could sit down during the breaks in between the blocks. In the experiment, participants watched a looped 8 min fragment of the animated film Animal Farm (Halas and Batchelor, 1954), played without sound (to minimize distraction), and projected on a screen in a city environment, akin to an open-air cinema (Figure 1).

Participants stood behind a group of 10 virtual agents who watched the movie together with them. The agents had a fixed position and were placed so that all 10 were clearly visible from the perspective of the participant. All trials started with a silent period of 1500–6500 ms (randomized in steps of 500 ms) in which nothing happened, followed by one of three sounds: an explosion, glass breaking, or a structure collapsing. The sounds were non-directional, with two of the sounds indicating that participants should look up either to the left or to the right (forced choice trials) and the third sound indicating that they could choose which side to look (free choice trials). The mapping of the sounds to the responses was counterbalanced. Together with the sound (Sample 1) or a random 100, 200, or 300 ms after the sound (Sample 2), a variable number of agents (0–10) also looked up to the left or right. During this time, the other agents kept on looking forward. The gaze direction was the same for all moving agents and was randomized with a 50% left, 50% right ratio across trials. As a result, on forced choice trials, participants had to look in the same direction as the agents in one-half of the trials (congruent trials) and in the opposite direction in the other half (incongruent trials). Each combination of sound cue, virtual agent gaze direction, and number of virtual agents looking up occurred once per block, in randomized order. The duration of the silent period at the start of each trial and the SOA were randomized across blocks, independently of the other variables.

Participants' task was to look at a fire presented in a window at the instructed or chosen location (Figure 1). When participants looked at the fire, it went out, signaling that their response was registered. Importantly, the fire was always presented at both locations, and participants were informed of this fact. This ensured that the virtual agents provided no relevant information about what participants would see, thereby allowing us to isolate reflexive, bottom-up processes. After looking at the fire, participants had to look back at the movie screen again. The next trial started when participants were looking at the screen and at least 2000 ms had passed since the agents moved their head. Participants were asked to respond as fast as possible but without making errors. In addition, they were asked to not use a strategy in the free choice trials but instead to choose randomly in the moment, so that they looked roughly equally often in both directions (e.g., Arrington and Logan, 2004).

After the experiment, participants completed a Dutch version of the Igroup Presence Questionnaire (IPQ; Schubert et al., 2001) as an exploratory measure of how participants perceived the virtual environment. The IPQ is a questionnaire designed to measure the subjective sense of being in a virtual environment. It includes a single item measuring the general sense of being present in the environment (*In the computer-generated world I had a sense of "being there"*; General Presence) and three multi-item subscales measuring the sense of being physically present in the environment (e.g., *I felt present in the virtual space*; Spatial Presence), the degree of being captivated by the virtual environment (e.g., *I was not aware of my real environment*; Involvement), and the realism of the virtual environment (e.g., *The virtual world seemed real to me*; Experienced Realism). All items were rated on a 7-point Likert scale.

QUANTIFICATION AND STATISTICAL ANALYSIS

The x, y, and z coordinates of the HMD were tracked continuously throughout the experiment. This allowed us to extract not only reaction times, error rates, and follow choices, but also movement times, partial errors, and partial choices. The algorithm to calculate these different dependent measures was developed on the pilot data mentioned earlier (Data S1). A visual depiction of how the algorithm works is provided in Figures S4 and S5. Reaction time was defined as the onset of upwards movement towards the target and was calculated by taking the first derivative of the HMD position on the y-axis and searching for the last point at which the first derivative was ≤ 0 before reaching its maximum. Because the first derivative reflects velocity, this method identifies the moment at which the response started. Movement time was defined as the time between the start and end of the movement and was calculated by searching for the first point at which the y-axis first derivative was ≤ 0 after reaching its maximum and then subtracting the reaction time from this time point. Note that reaction and movement times were based on the HMD position on the y- and not z-axis because the z-axis movement onset was often ambiguous (i.e., participants tended to move their head more left/right than up/down in between trials). However, using the z- instead of y-axis positions yielded very similar results (Data S3).

Partial errors (forced choice trials) and partial choices (free choice trials) were defined as a substantial deviation in the direction opposite to the eventually chosen direction and were calculated by taking the first derivative of the HMD position on the z-axis, coding it so that positive values always reflected a movement in the direction of the eventually chosen target, and then searching for a five-point local minimum of < -0.05 preceding the maximum first derivative value. The chosen response and its accuracy were defined based on the first target participants hit. On free choice trials, the chosen response was coded as a "follow choice" or a "not follow choice" depending on whether participants followed the gaze of the virtual agents.

Exclusion criteria were also based on the aforementioned pilot study (Data S1) and were preregistered (https://aspredicted.org/blind.php?x=cu9ma6). The following criteria were used to exclude participants. First, participants were excluded from all analyses if their error rate on forced choice trials was $\geq$40% (0 participants). Second, participants were excluded from the forced choice analysis if their error rate was $\geq$2.5 SD above the sample's mean error rate (6 participants) or if their reaction time on forced choice trials was $\geq$2.5 SD above the sample's mean reaction time on those trials (3 participants). Finally, participants were excluded from the free choice analysis if they looked in the same direction on $\geq$ 75% of the trials (12 participants) or if their reaction time on free choice trials was $\geq$2.5 SD above the sample's mean reaction time on those trials (2 participants). As a result, 151 participants were retained for the forced choice analysis and 146 participants for the free choice analysis.

The following criteria were used to exclude trials. First, trials in which none of the agents moved were excluded from all analyses because congruency and follow choice were undefined on those trials. These trials were included only to discourage a response strategy in which participants used the agents' movements as a cue to respond. The mean reaction time on trials where the agents did not move was 612 ms for forced choice trials and 682 ms for free choice trials. Second, trials were excluded from all analyses if the reaction time was $\leq$200 ms (0.73%) or $\geq$4000 ms (0.01%), or if the movement time was $\geq$2000 ms (0.01%). Third, trials were excluded from the forced choice reaction and movement time analyses if the response was incorrect (5.17%), if a partial error was identified (5.18%), or if the reaction (1.71%) or movement time (1.91%) was $\geq$2.5 SD above or below the mean reaction or movement time on forced choice trials. Fourth, the same trials were also excluded from the forced choice partial error analysis, except that partial errors were included, and from the forced choice error rate analysis, except that errors were included. Fifth, trials were excluded from the free choice reaction time, movement time, and follow choice analyses if a partial choice was identified (3.04%) or if the reaction (1.56%) or movement time (1.89%) was $\geq$2.5 SD above or below the mean reaction or movement time on free choice trials. Finally, the same trials were also excluded from the free choice partial choice analysis, except that partial choices were included.

Reaction and movement times were analyzed with linear mixed effects models and error rates, partial errors, partial choices, and follow choices were analyzed with generalized linear mixed effects models using a binomial logit link function (Baayen et al., 2008; Bates et al., 2014). All models included group size (i.e., the number of virtual agents making a movement) as a centered trial-by-trial numerical predictor. The forced choice models further included congruency (congruent or incongruent) as a factor. The free choice reaction time, movement time, and partial choice models instead included the participant's choice (follow or not) as a factor. Although we had preregistered to also include SOA, we eventually did not do this for two reasons. First, in the combined data, SOA was partly confounded with sample. Second, not all models converged when SOA was included. However, including SOA in the analysis of Sample 2 when possible had very little influence on the results (Data S4). The random effects structure of the models was determined using a backwards selection procedure because this has shown to optimally balance Type I and II error rates (Matuschek et al., 2017). p-values for the linear mixed effects models were calculated using *t* tests with Satterthwaite-corrected degrees of freedom (Kuznetsova et al., 2017) and p values for the generalized linear mixed effects models were calculated using Wald tests. The primary dependent variable for the forced choice analysis was reaction time. The primary dependent variables for the free choice analysis were reaction time and follow choices. Importantly, whereas reaction times are a rather indirect measure of gaze following, follow choices provide a more direct gaze following measure, similar to how it is performed in real life. Movement times, partial errors, and errors were analyzed as secondary dependent variables.

Finally, as a preregistered exploratory analysis, we also looked for non-linear patterns in the data, based on the fact that previous research has mostly found asymptotic relationships between group size and gaze following (Gallup et al., 2012; Jorjafki et al., 2018; Milgram et al., 1969). Non-linear models were fitted by replacing the linear group size term with a linear b-spline term with a single knot at the average group size (Hastie et al., 2009). This approach effectively fits two lines: one between group sizes 1 and 5 and one between group sizes 6 and 10. As such, it allows us to test if the influence of group size on behavior reaches an asymptote.