Article

Social Psychological and Personality Science
1-10

DOI: 10.1177/1948550617709112
journals.sagepub.com/home/spp
SAGE

Strangers in a Stadium: Studying Group Dynamics With In Vivo Behavioral Tracking

Joshua Conrad Jackson[1], David Bilkey[2], Jonathan Jong[3], Maya Rossignac-Milon[4], and Jamin Halberstadt[2]

Abstract

Social group dynamics are a defining topic of psychological science, yet the field still lacks methods of tracking groups with precision and control. Previous methods have been hampered by limitations either to external validity (e.g., ecologically deficient environments) or to internal validity (e.g., quasi-experimental designs), but a new technique—which we term in vivo behavioral tracking (IBT)—resolves this trade-off. Through IBT, we track large numbers of people in controlled environments over time, while storing precise behavioral data that can be linked to information regarding participants' attitudes, personality, and demographics. In this article, we describe the fundamentals, assumptions, and challenges of IBT methodology. We also compare IBT to other tracking methods and illustrate some insights it has provided into group formation and cooperation. We argue that IBT is a highly valid and surprisingly feasible method of studying groups that should be used alongside more traditional forms of data collection.

Keywords

in vivo behavioral tracking, group behavior, cooperation, group formation, social dynamics, social psychology

Henri Tajfel argued that studying group behavior was "the most urgent and ominous task confronting us at present" (Tajfel, 1981, p. 128). Tajfel's claim remains as true today as it was 25 years ago. Our species' most memorable moments—from wars, to festivals, to revolutions—have been defined by human behavior in groups, yet we still remain largely unaware of how the mind operates in large groups or how groups can seemingly take on minds of their own. Social psychology's contributions to understanding group behavior have involved rigorous empirical analysis. With creative laboratory paradigms and field studies, social psychologists have shed light on hundreds of phenomena, giving us insight into how social pressures influence attitudes (Bem, 1967; Festinger, 1962; Latané, 1981), behaviors (Cialdini, Reno, & Kallgren, 1990; Milgram, 1974), and even basic cognitive processes like memory and attention (Baldwin, Bagust, Docherty, Browman, & Jackson, 2014; Coman, Manier, & Hirst, 2009; Shteynberg, 2010).

These contributions notwithstanding, many of our field's paradigms suffer from a methodological trade-off between external and internal validity, which prevents us from studying large groups with the same degree of control that characterizes effective research on individual behavior. In this article, we introduce *in vivo behavioral tracking* (IBT) as a new method that can resolve this trade-off. In our variant of IBT, individuals interact in a fully enclosed stadium as their behavior is tracked surreptitiously from a camera mounted high and directly above them. The resulting data offer insight into group formation, the emergence of group norms, and other forms of dynamic social behavior, with high experimental control and ecological validity. While IBT may sound prohibitively expensive or logistically complex, we show how the method's flexibility makes it accessible and feasible to implement.

Social Psychology's Methodological Trade-Off

Social psychologists study intergroup and cognitive processes with a wide and growing array of methods, which vary in their internal and external validity. At one extreme are unobtrusive observations of group behavior in natural environments. These have evolved from ethnographic methods in anthropology but are typically more hypothesis driven in their psychological

[1] Department of Psychology, University of North Carolina, Chapel Hill, NC, USA
[2] Department of Psychology, University of Otago, Dunedin, Otago, New Zealand
[3] Centre for Research in Psychology, Behaviour, and Achievement, Coventry University, Coventry, UK
[4] Department of Psychology, Columbia University, New York, NY, USA

Corresponding Author:
Joshua Conrad Jackson, Department of Psychology, University of North Carolina, Chapel Hill, NC 27599, USA.
Email: joshcj@live.unc.edu

instantiations. In one classic example, Freeman and Webster (1994) observed that beachgoers' physical locations over 31 days could not only predict their interaction frequency but also their inferences of psychological similarity. A more recent example is "free range data harvesting," in which the similarity among randomly sampled dyads is used to infer the "fault lines" around which social groups emerge, and the contextual moderators of dyadic similarity (Crandall, Schiffhauer, & Harvey, 1997). In one such investigation, Bahns, Pickett, and Crandall (2011) found that dyads tended to have more in common at large universities compared to small universities. The authors speculatively attributed the effects to greater relational mobility at large campuses, but they could not experimentally test this account, just as Freeman and Webster could not directly test the facets of proximity that accounted for its relationship with affiliation. Both studies suffer from the same limitation, seemingly unavoidable in observational field studies: The proposed psychological processes underlying observed naturalistic behavior can neither be manipulated nor precisely measured.

At the other end of the spectrum are laboratory studies designed to maximize environmental control and internal validity. Early intergroup laboratory paradigms included the bogus stranger paradigm and the minimal group paradigm. In the former method, participants rate their attraction to a fictitious stranger based on cursory knowledge of the stranger's preferences (e.g., Byrne, Clore, & Worchel, 1966). In the latter, participants are assigned group membership based on trivial characteristics, and then make relevant intergroup judgments such as choosing how to distribute shared resources (e.g., Turner, 1982). Both paradigms have yielded insights into intergroup bias and homophily and have been complemented by newer and more precise social cognitive measurements of implicit bias, such as the Implicit Association Test (Greenwald, McGee, & Schwartz, 1998; Greenwald, Nosek, & Banaji, 2003) and the Affective Misattribution Procedure (Fazio, 2001; Payne, Cheng, Govorun, & Stewart, 2005).

Computational methods have recently emerged as an even more precise way of experimentally studying group behavior by simulating environments of many individuals. For example, agent-based models sample artificial agents, which represent intentionally simplified humans behaving in theoretically consistent ways (Jackson, Rand, Lewis, Norton, & Gray, 2017; Macy & Flache, 2009). Since agent-based models are computer simulations, researchers can use them to create and control environments, samples, and modes of interaction. Agent-based models also allow researchers to study dynamic phenomena, which emerge over many generations through a transactive relationship between individuals and their environments (Smith & Conrey, 2007).

Nevertheless, these controlled methods are, almost by definition, simplifications of the environments they hope to model, requiring additional assumptions to make inferences from their results. Participants' self-reports of intergroup attitudes are often inaccurate (e.g., Dovidio, Gaertner, Kawakami, & Hodson, 2002; Holmes, 2009), and people's decisions in economic games often do not translate to other forms of prosociality (Bardsley, 2008; Levitt & List, 2007; Winking & Mizer, 2013). Agent-based approaches, despite being a promising new method in social psychology, are only as valid as the models and assumptions that drive their algorithms and sometimes must be replicated with human participants to verify that they behave as predicted.

Other methods alternatively have more internal or external validity but come with weaknesses of their own. Archival analyses can track attitudes in large groups but cannot experimentally manipulate the environments in which those attitudes are measured. Field experiments can improve on the laboratory's external validity but tend only to manipulate conditions among dyads and small groups of individuals and often lack a priori information about their samples. In fact, all methods appear to have at least one of the two general limitations. Researchers lack the scale, resolution, or control to quantify the process behind their observations or they lack the ecological validity to apply precise quantitative findings to real-world phenomena.

Resolving the Trade-Off Through IBT

IBT draws from automated image-based tracking research in animal behavior, wherein data on movement and position are converted from tracking preassigned visual objects (e.g., patches of color) over the course of a video record (see Dell et al., 2014). This mode of tracking can provide precise data but has (until now) been largely used in correlational designs. There is some limited human research that uses environmental manipulations with image-based tracking methodology (e.g., Gallup et al., 2012; Zhan, Monekosso, Remagnino, Velastin, & Xu, 2008), but these too are only quasi-experimental and also lack information about participants' demographic and psychological characteristics. The age and gender of individuals on a city street can be roughly approximated, but their personality traits, self-esteem, and interpersonal preferences remain hidden and unknown.

IBT moves beyond previous human and nonhuman tracking paradigms by combining environmental control, precise measurement, ecological validity, and rich information on individual differences. Participants—who have previously provided relevant demographic and individual difference data—are unobtrusively filmed (in our case, by a small camera 25 m above them) in a controlled environment large enough to permit relatively unconstrained behavior. Participants' position in the video feed can then be converted into any number of spatial or movement data (e.g., physical location, speed, proximity), which can be merged with pretest data containing their attitudes and demographic information. Thus, the methodology produces precise real-time data, which can be analyzed as a function of any measurable individual difference. Furthermore, the measurement technique can be combined with nearly any laboratory manipulation by exposing participants to different conditions before they enter the experimental venue.

The general steps involved in implementing an IBT study are as follows.

Step 1: Find a Suitable Venue to Mount a High-Resolution Camera

To track participants' in vivo behavior, one needs a clear view of all participants throughout the study. Otherwise, IBT is compatible with areas of variable size and layout. In our studies, we have used a section of the Forsyth Barr stadium in Dunedin, New Zealand. The venue is fully enclosed, allowing us to hang a mounted camera on a beam 25 m above the stadium's floor, sufficiently high to be practically invisible and to maintain a clear view of all participants throughout the experiment. To track behavior, our group has used an Elphel NC535 network camera, which records at 30 frames/second at a full resolution of 2,592 $\times$ 1,944 pixels. A Theia SY110 lens provides a 120° view with almost 0% distortion, which ensures precise error-free tracking.

Step 2: Recruit a Sample and Administer Premeasures

Before experimentation, researchers using IBT must recruit a sample and gather theoretically relevant self-report and demographic information. For researchers studying the dynamics of small group behavior, the IBT paradigm is applicable to samples of fewer than a dozen individuals. For those studying large crowds or swarm behavior, the IBT approach is conceivably applicable to samples of hundreds or even thousands of participants. In our own studies, we have sampled multiple groups of approximately 50 participants, who have completed relevant psychological measures online several weeks prior to experimentation.

Step 3: Run Behavioral Activities Within the Experimental Venue

The IBT parameters allow for a huge range of behavioral activities to test any number of research questions. However, variables that can be operationalized in terms of changes in physical position over time (e.g., velocity, acceleration, trajectory) take better advantage of the unique power of tracking data. Our own experimental tasks have been as simple as asking participants to follow a research assistant around the venue (see Figure 1) or "get into groups" (from which groups' speed of formation, composition, and duration can be derived and linked to participants' individual differences) and as complex as a large-scale "foraging" task meant to approximate group survival activities in preagricultural societies. Other researchers will, of course, have other priorities.

Needless to say, IBT has a better claim to ecological validity when participants are unaware they are being filmed, which may pose an ethical issue for some review boards. In our case, we have reached a compromise position: Participants in our studies consent to appear on video but are not told (nor has anyone inquired about) the specifics of how, when, or where any filming would take place. Participants indicate no awareness of the mounted camera. After the study, participants are e-mailed a fuller debriefing, which includes the option to remove their data from analysis. No participants have taken up this option.

Step 4: Generate Participants' Positional Data

After collecting data, researchers must extract and quantify participants' physical location over the course of the study. For us, a key challenge in this step is determining who is who, since it is nearly impossible to confirm participants' identities from our video feed alone. We address this problem by having participants enter the experimental area one by one at the beginning of the study, in order of identification codes they have been randomly assigned before the study. Participants' identification numbers are linked to their questionnaire responses, allowing us to reliably match the two sets of data.

Once participants' numbers are identified in the video feed, computer software is required to follow them automatically as they move about the experimental space. Our specific tracking software—developed by the software engineering group Animation Research Limited—extracts sets of image patches for each participant and stores x–y coordinates associated with these patches. It updates patch coordinates in each subsequent image of the video sequence using computer vision methods, such as template matching (Lewis, 1995) and histogram-based matching (Porikli, 2005). To use the software, the user first selects a tracking target of interest (i.e., a participant in the study) with a mouse-click on the target's head, which defines a unique template (a box around the starting point) to be matched (see Figure 2 for a view of these templates linked with their participants). The program assumes that all participants are part of the dynamic foreground. It then creates a template of extracted foreground pixels that will be traced for each frame of the video, calculating the location of the best match between each new frame's template and the template that was previously stored. In our research, we used a camera that recorded at 30 frames/second, generating 30 sets of coordinates per second per participant.

Any visual tracking method—including ours—faces two challenges in trying to distinguish a participant from his surrounding environment. The first is occlusion; participants may be "lost" or misidentified as they pass behind one another. The number of participants, the nature of their interaction, and the angle at which they are filmed all influence the likelihood of occlusion. Although the former two factors may be constrained by the participants and venue to which researchers have access, the latter can be addressed by filming from as steep an angle as possible, and ideally directly overhead, as we did in the work reported here.

The second challenge is that participants may be lost against the background, a problem that can be minimized by filming in good light and on a clean and bright surface. In our method, we took the additional step of supplying participants with orange hats to maximize contrast with the stadium's cement floor. Such interventions risk compromising ecological validity (e.g., in the case of hats, by making participants uncomfortable or suspicious) so they should be used with caution and with an appropriate cover story. Furthermore, even after taking such measures, researchers should not assume any automated tracking will be error free. In our case, a research assistant monitors

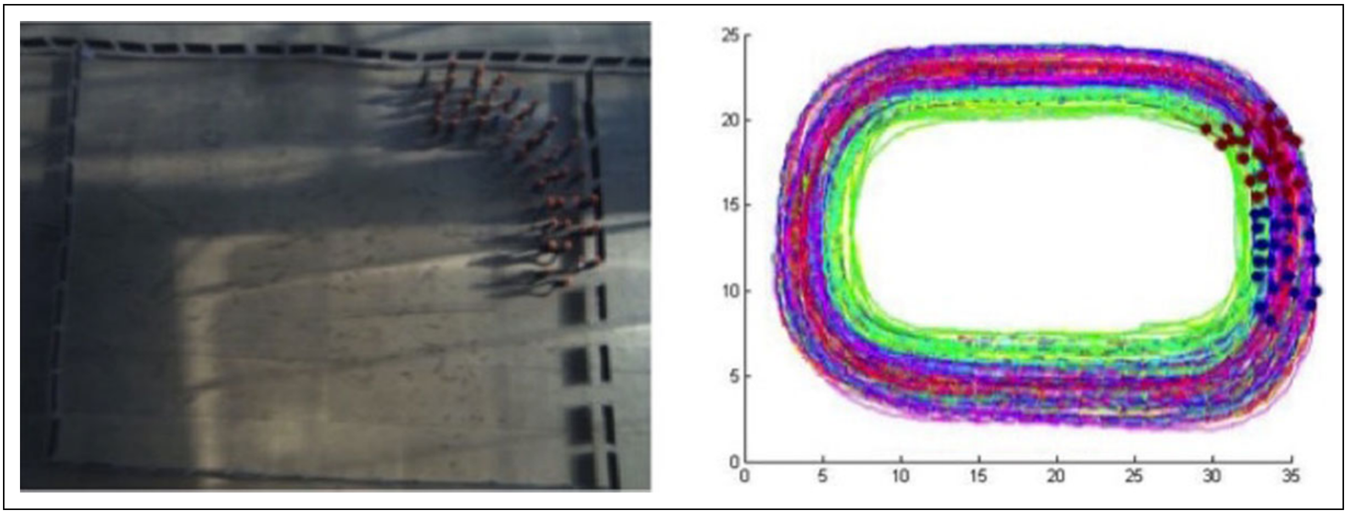

Figure 1. An overhead view of participants following a research assistant around the experimental venue, complemented with a view of the tracking data as they are later processed and analyzed via a custom MATLAB script. Participants are shown in slightly different positions across these two images, since the images represent slightly different time points. Colors represent the path of different individuals. Scale is in meters.

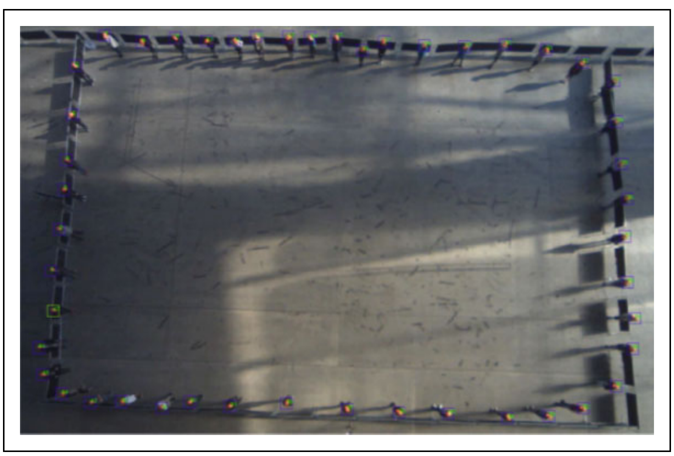

Figure 2. An image of the video feed as participants' locations is being tracked. Boxes over participants' heads represent color templates that are matched and updated over the course of the experiment, resulting in sequential *x–y* coordinates.

the software's progress and makes manual adjustments on the rare occasions that a tracking patch becomes displaced.

Step 5: Analyze Tracking Data and Integrate With Other Data

In order to analyze behavioral data in relation to premeasures, one must use participants' *x–y* coordinates throughout the study to compute speed and interindividual proximity indices. In our case, we have created a custom MATLAB Version 8.3 script that, for any user-defined interval, generates measures of average speed and distance traveled, instantaneous speed at the conclusion of that interval, and measures of interpersonal proximity. It also generates more specific information regarding participants' clustering using a k-means algorithm followed by silhouetting, which quantitatively identifies discrete social groups based on the ratio of participants' proximity to group members versus nongroup members (see Figure 3). The program returns fit coefficients for groups of different sizes, which can be used to determine the most accurate number of groups that formed during the prespecified interval, along with the members of each group and their positions within it. Finally, all movement and group membership data are merged with individual difference or experimental data collected prior to the study.

Assumptions of IBT

The primary advantage of IBT is its ability to measure behavior unobtrusively with high precision and control. Nevertheless, the method requires a few basic assumptions.

Proximity as Affiliation

IBT's first assumption is that proximity is a proxy for affiliation and that any convergence of individuals is nonrandom and psychologically significant. This assumption draws from the classic literature on personal space, which delineates a proximal boundary into which only close others are permitted (Hayduk, 1983), such that approaching that boundary increases psychological intimacy. We argue that when observed on a large scale, patterns of physical proximity yield insights into emergent social ties.

Speed as Effort

In many important group activities, from foraging, to hunting, to warfare, speed relates to group commitment, and we have made use of this association in our IBT designs. In one study,

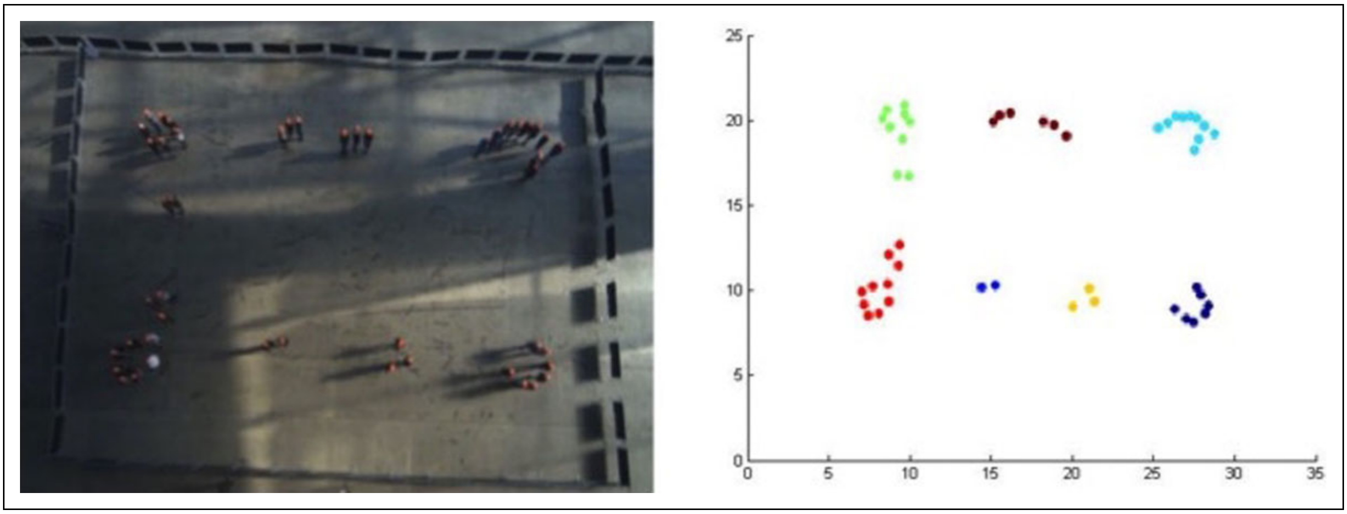

Figure 3. An overhead view of participants completing a group formation task in which they were asked to "get into groups of any size or composition" over several iterations. Using a MATLAB k-means/silhouetting algorithm, we were able to quantitatively identify social groups via participants' spatial distribution and extract data about these groups. The right panel displays these quantitatively derived groups. Scale is in meters.

we asked participants to collect hundreds of tokens scattered across the experimental area, with instructions that the experiment would not end until all tokens had been collected. Since it is in everyone's interest to collect the tokens promptly, but in no individual person's interest to be the one to collect them, this situation represents a large-scale behavioral social dilemma, and the speed with which participants search for tokens reflects their willingness to cooperate for the good of the group at the expense of their own effort. Speed can also offer insight into more nuanced social motivations. For instance, in a marching task, an individual's synchronization with his or her group's walking speed might reflect their desire to conform to descriptive norms, while deviating from the group's speed might indicate low motivation to socially coordinate.

IBT Approximates the Field on Psychologically Relevant Dimensions

By definition, no controlled space is exactly like real life. Methodology-specific reactivity can be quantified (e.g., Mehl & Holleran, 2007), and IBT researchers may wish to do so, particularly if they reuse a particular venue over a series of studies. However, an equally relevant issue is whether a potentially unique experimental context mimics real life in theoretically relevant ways. In the case of large group behavior, the theoretically relevant ways include participants' ability to naturally form groups and cooperate with each other without the interference of an experimenter or the confines of a closed laboratory space. Thus, although a stadium (or any other IBT venue) hardly resembles a battlefield, church, or urban sprawl, it remains an appropriate venue to experimentally study the group behaviors that typify these spaces.

How Does IBT Compare With Other Tracking Methods?

IBT is not the first tracking method that has been applied to human behavior. There are several other tracking technologies available with the potential to study behavior in large groups. These technologies differ in terms of their precision, operational constraints, and costs.

Global Positioning System (GPS)

The cost and size of GPS receivers have decreased markedly and there has been a corresponding increase in their use to track nonhuman species ranging from whales (Wahlberg, 2002), to bats (Tsoar et al., 2011). Furthermore, as these receivers have become ubiquitous, participants in a human study could feasibly carry GPS tracking devices (indeed, most people are already carrying them in their smartphones), with their *x*–*y* (and *z*) data continuously streamed to a central location or downloaded offline.

There are, however, two major limitations that hamper the research application of GPS software. The first of these relates to spatial resolution: GPS provides a 95% accuracy of around 3–4 m, but this can be poorer under some circumstances, as when there is an ionospheric delay (which interferes with the satellite strength as it penetrates the earth's atmosphere; Sardon, Rius, & Zarraoa, 1994). GPS error can be reduced to as little as 1 m when coupled with other technologies, such as the Wide Area Augmentation System (Parkinson & Spilker, 1996), which is available in limited geographical locations and devices. This resolution, however, may still be insufficient to detect the subtle changes in proximity that reflect social affiliation.

Ultra-Wideband (UWB)

Alternative tracking technologies have been developed based on radio transmitters such as the radio frequency identification and UWB. These systems consist of one or more fixed and calibrated receivers and mobile tags worn by study participants. For UWB, the tags would transmit UWB radio pulses to linked sensors arranged around an experimental area. These sensors use time difference of arrival and angle of arrival data to determine transmitter location. Although some of these devices have very good sampling rates, they require a calibrated sensor infrastructure that may be challenging to set up for short-term studies. Furthermore, as many of these devices were developed for commercial applications such as inventory management, it is unclear how they perform with a large number of sensors moving rapidly through the area of interest, or in cases of high signal occlusion due to large samples.

Social Sensors

Social sensors include much of the functionality of UWB tags but do away with the complex infrastructure. Early social sensors were designed for a variety of purposes, such as capturing teacher–student interactions, and triggering automatic doors based on the wearer's position (Borovoy, McDonald, Martin, & Resnick, 1996; Olguín & Pentland, 2008; Want, Hopper, Falcao, & Gibbons, 1992). More recent innovations have included the *sociometric badge*, which are able to capture participants' physiological states and interpersonal behavior (including orientation to other participants; Choudhury & Sabherwal, 2003; Olguín, Paradiso, & Pentland, 2006). They also recognize common daily activities (e.g., sitting, running) in real time with at least 80% accuracy (Olguín & Pentland, 2006) and can analyze wearers' speech patterns for interest and excitement (Pentland, 2005).

However, sociometric badges have been largely restricted to use in organizational contexts, such as analyzing staff behavior in hospitals (see Rosen, Dietz, Yang, Priebe, & Pronovost, 2014), though some exploratory work has used sociometric badges to analyze face-to-face interactions during coffee breaks (Atzmeuller, Ernst, Krebs, Scholz, & Stumme, 2014) and gender differences in cooperation (Onnela, Waber, Pentland, Schnorf, & Lazer, 2014). Their limited use is partially due to the cost of the badges but also due to their relative imprecision. Badge-derived spatial data have significant noise, with accuracy ranging from 1 to 3 m in previously published research (Cattuto et al., 2010; Onnela et al., 2014).

In sum, there are several existing and emerging options for tracking participants without a mounted camera. Some (e.g., GPS) are relatively inexpensive but are too imprecise to approximate movement and speed with the appropriate resolution and scale. Others are more precise but are not feasible in most contexts due to a limited infrastructure (e.g., UWB methods) or to their high cost (e.g., sociometric badges). Measures do exist that can precisely gather data on participants' conversations and emotional states (e.g., electronically activated recorder; Mehl, Pennebaker, Crow, Dabbs, & Price, 2001), yet no spatial tracking alternative can rival IBT's combination of precision, contextual flexibility, and affordability.

Two Questions Answered With IBT

To illustrate the method's strengths, we next explain our own applications of IBT to two fundamental questions in social psychology: How do people form social groups? And when do people cooperate with their groups? Our hypotheses were based on preexisting social science literature but had not been definitively tested given the aforementioned constraints of traditional laboratory paradigms.

How Do People Form Social Groups?

Some of the earliest research on group formation came from social identity theorists, who argued that individuals identify with and favor others who share common features and that groups are likely to form on the basis of such features (Brewer & Kramer, 1985; Tajfel, 1982). These researchers found, for example, that participants allocated relatively more resources to others with whom they shared group membership, even if that affiliation was arbitrary and largely meaningless (e.g., based on preference for abstract art or nametag color; Billig & Tajfel, 1973; Tajfel, 1978, 2010).

Yet many of these early paradigms suffered from considerable limitations. Most notably, people in these studies typically made decisions alone, working in cubicles with digital or paper forms. Even when participants believed that they were interacting with group members, these interactions were often simulated by the experimenter and did not involve other people (e.g., Ellemers, Kortekaas, & Ouwerkerk, 1999). This limitation means that many findings in the field of social identity are based on mere approximations of real group situations. A second limitation of previous designs was their focus on single interactions—a trend that continues with the heavy reliance of social psychological research on one-shot dilemmas (Camerer & Fehr, 2006; Smith & Conrey, 2007). Single-shot paradigms ignore the dynamic nature of social groups, wherein processes like homophily and ostracism can snowball or diminish over time.

To try to address these limitations, we applied IBT to the question of group formation (Halberstadt et al., 2016). We positioned samples of approximately 50 individuals around our experimental area and repeatedly instructed them to form groups of any size or composition. When we linked participants' location data with their demographics, we found that group formation occurred primarily on the basis of participants' physical attractiveness (as rated by independent coders after the study) and their gender. However, both effects decreased over trials, with groups getting more heterogeneous over time. This evidence has points of both convergence and divergence with previous research. As in previous studies using more restricted paradigms, participants preferred to interact with others who had salient physical features in common with

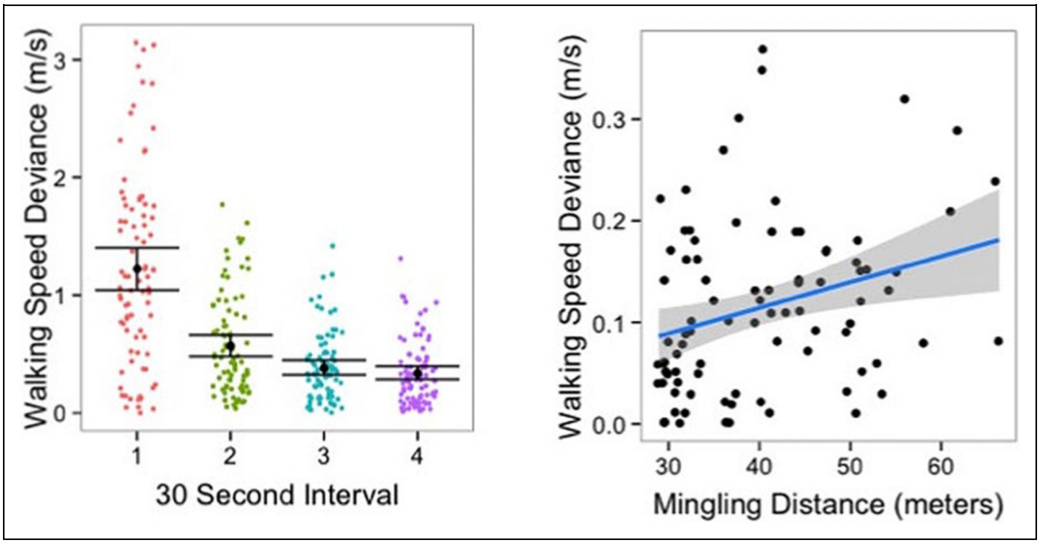

Figure 4. Synchronization of group walking speed over time. During a follow-the-leader walking task, deviance from the average group walking speed decreased, such that participants progressively synchronized their speed over time. Participants who tended to synchronize their speed with the group speed also later stood in greater proximity to their peers, suggesting that the tendency to synchronize and affiliate with others may be linked. In both panels, walking speed deviance is represented as meters per second. Since the left panel is distributed across time-points, it contains more extreme values.

them. However, in contrast to other research on social grouping (e.g., Gray et al., 2014; Schelling, 1971), groups did not become more similar with time. Rather, they became more diverse, suggesting that superficial physical cues might be weighted less as group members become more familiar.

In another study, we added an experimental component to our group formation task, giving each participant one of the two different colored nametags in a real-life version of the minimal-group paradigm, while also measuring self-esteem and collective self-esteem. Groups segregated themselves by nametag color to a greater extent than expected by chance, demonstrating for the first time that the minimal group manipulation actually produces *groups* (rather than just bias). Furthermore, participants' individual and collective self-esteem predicted the minimal group effect in opposite ways. Those with high individual self-esteem showed the weakest tendency to group by nametag color, while participants with high collective self-esteem showed the strongest, suggesting that minimal grouping is differentially valuable depending on the level at which people derive their esteem.

In addition to these homophily effects, IBT has yielded insights into grouping heterophily—the process whereby individuals *avoid* grouping with others who share common traits. Previous research has suggested that individuals who are high in anxious attachment might prioritize warm partners in their social relationships (e.g., Feeney & Noller, 1990), while those high in avoidant attachment might prioritize partners who offer them autonomy (e.g., Mayseless & Scharf, 2007). Yet, while avoidantly attached individuals might be well equipped to provide the relational autonomy that they desire in others, anxiously attached individuals tend to display high social anxiety (Cash, Theriault, & Annis, 2004) and distrust (Knobloch, Solomon, & Cruz, 2001) in their relationships, making them ill-suited as partners for other anxiously attached individuals. Our IBT data showed grouping effects consistent with this possibility: People high in attachment avoidance tended to form groups with other avoidantly attached individuals, but those high in attachment anxiety tended to form groups with others who were *low* in attachment anxiety. These data suggest that the behavioral cues associated with attachment styles (see McClure & Lydon, 2014; Shaver, Schachner, & Mikulincer, 2005) might be detectable in brief nonverbal interactions and appear to shape early group formation amongst strangers.

When Do We Cooperate With Our Groups?

The origin of large-scale human cooperation is one of the most intriguing issues in the social sciences, occupying a significant portion of the literatures in economics (Axelrod, 1980; Hamilton & Axelrod, 1981), sociology (Durkheim, 1893), behavioral ecology (Boyd & Richerson, 1988), anthropology (Whitehouse, 2012), and psychology (Bear & Rand, 2016; Norenzayan et al., 2016). Many of these disciplines conceptualize cooperation as the driving force behind the proliferation of large human groups, but also see it as counterintuitive, since people live in large anonymous communities of nonrelatives where defection is often a more favorable strategy for resource accrual and individual survival. Indeed, in our IBT studies, we have found that individuals who were more deeply embedded in their groups early in the experiment also put less effort into a subsequent cooperative task, suggesting that anonymity is associated with defection in group tasks.

We have also used IBT to explore one proposed solution to the problem of cooperation: the development of kin-like ties through ritualized behavior (Durkheim, 1915; Whitehouse & Lanman, 2014). Experimental investigations had previously found that common elements of rituals such as behavioral synchrony (Wiltermuth & Heath, 2009) and shared pain (Bastian, Jetten, & Ferris, 2014) independently increased group commitment in social dilemmas. However, their generalizability was limited due to the small size of their groups: These studies had exclusively examined groups of fewer than five members, even though most rituals involve much larger collectives (Fischer, Callander, Reddish, & Bulbulia, 2013). Using IBT, we sought to extend the previous work to larger groups, while testing how different elements of rituals might interact. We designed a pseudoritual that masqueraded as a follow-the-leader task; participants had to follow a research assistant around the experimental area for several minutes. We found, first, that participants as a group synchronized their walking speed over time (see Figure 4) and that participants who did so more fully stood in closer proximity to their peers at later points in the experiment, indicating a potential link between synchrony and prosociality in large groups. In an experimental design, we also systematically varied participants' walking synchrony (via instructions to walk in-step with other participants) and arousal (via walking speed). Tracking data revealed that participants who moved synchronously (vs. asynchronously), and those who moved quickly (vs. at a normal pace), later formed larger groups, stayed closer to the members of those groups, and were more cooperative in the "foraging" task described above. Critically, the effects of synchrony and arousal were larger in combination than alone.

Conclusion

Social psychology is often defined as the study of how humans behave in groups, yet many of our field's paradigms are not well suited for studying naturalistic group behavior. Existing methodological limitations can either be traced to issues with internal validity (e.g., quasi-experimental designs and imprecise measurement) or with external validity (e.g., ecologically deficient environments or poor operationalization). In this article, we showed how IBT can help resolve this trade-off, reviewed some of the practical and methodological factors to consider when using the paradigm, and illustrated some initial applications to basic questions of group dynamics. We conclude that IBT is an important counterpart to traditional and emerging laboratory and computational paradigms in understanding how groups form and evolve over time.

Declaration of Conflicting Interests

The author(s) declared no potential conflicts of interest with respect to the research, authorship, and/or publication of this article.

Funding

The author(s) disclosed receipt of the following financial support for the research, authorship, and/or publication of this article: The work in this paper was supported by a Large Grant from the UK's Economic and Social Research Council (REF RES-060-25-0085) entitled "Ritual, Community, and Conflict" and an award from the John Templeton Foundation entitled "Religion's Impact on Human Life," and an Advanced Grant from the European Research Council (ERC) under the European Union's Horizon 2020 research and innovation programme (grant agreement No. 694986).

References

Atzmueller, M., Ernst, A., Krebs, F., Scholz, C., & Stumme, G. (2014, April). On the evolution of social groups during coffee breaks. In C. Chung (ed.), *Proceedings of the 23rd International Conference on World Wide Web* (pp. 631–636). Seoul, Republic of Korea: ACM.

Axelrod, R. (1980). Effective choice in the prisoner's dilemma. *Journal of Conflict Resolution*, *24*, 3–25.

Bahns, A. J., Pickett, K. M., & Crandall, C. S. (2011). Social ecology of similarity: Big schools, small schools and social relationships. *Group Processes & Intergroup Relations*, *15*, 119–131.

Baldwin, M. W., Bagust, J., Docherty, S., Browman, A. S., & Jackson, J. C. (2014). The "Rodand Fran Test": Relationship priming influences cognitive-perceptual performance. *The Journal of Social Psychology*, *154*, 441–446.

Bardsley, N. (2008). Dictator game giving: altruism or artefact? *Experimental Economics*, *11*, 122–133.

Bastian, B., Jetten, J., & Ferris, L. J. (2014). Pain as social glue shared pain increases cooperation. *Psychological Science*, *25*, 2079–2085.

Bear, A., & Rand, D. G. (2016). Intuition, deliberation, and the evolution of cooperation. *Proceedings of the National Academy of Sciences*, *113*, 936–941.

Bem, D. J. (1967). Self-perception: An alternative interpretation of cognitive dissonance phenomena. *Psychological Review*, *74*, 183–200.

Billig, M., & Tajfel, H. (1973). Social categorization and similarity in intergroup behaviour. *European Journal of Social Psychology*, *3*, 27–52.

Borovoy, R., McDonald, M., Martin, F., & Resnick, M. (1996). Things that blink: Computationally augmented name tags. *IBM Systems Journal*, *35*, 488–495.

Boyd, R., & Richerson, P. J. (1988). *Culture and the evolutionary process*. Chicago, IL: University of Chicago Press.

Brewer, M. B., & Kramer, R. M. (1985). The psychology of intergroup attitudes and behavior. *Annual Review of Psychology*, *36*, 219–243.

Byrne, D., Clore, Jr. G. L., & Worchel, P. (1966). Effect of economic similarity-dissimilarity on interpersonal attraction. *Journal of Personality and Social Psychology*, *4*, 220.

Camerer, C. F., & Fehr, E. (2006). When does "economic man" dominate social behavior? *Science*, *311*, 47–52.

Cash, T. F., Thériault, J., & Annis, N. M. (2004). Body image in an interpersonal context: Adult attachment, fear of intimacy and social anxiety. *Journal of Social and Clinical Psychology*, *23*, 89–103.

Cattuto, C., Van den Broeck, W., Barrat, A., Colizza, V., Pinton, J. F., & Vespignani, A. (2010). Dynamics of person-to-person interactions from distributed RFID sensor networks. *PloS One*, *5*, e11596.

Choudhury, V., & Sabherwal, R. (2003). Portfolios of control in outsourced software development projects. *Information Systems Research*, *14*, 291–314.

Cialdini, R. B., Reno, R. R., & Kallgren, C. A. (1990). A focus theory of normative conduct: Recycling the concept of norms to reduce littering in public places. *Journal of Personality and Social Psychology*, *58*, 1015–1026.

Coman, A., Manier, D., & Hirst, W. (2009). Forgetting the unforgettable through conversation socially shared retrieval-induced forgetting of September 11 memories. *Psychological Science*, *20*, 627–633.

Crandall, C. S., Schiffhauer, K. L., & Harvey, R. (1997). Friendship pair similarity as a measure of group value. *Group Dynamics: Theory, Research, and Practice*, *1*, 133–143.

Dell, A. I., Bender, J. A., Branson, K., Couzin, I. D., de Polavieja, G. G., Noldus, L. P., . . . Brose, U. (2014). Automated image-based tracking and its application in ecology. *Trends in Ecology & Evolution*, *29*, 417–428.

Dovidio, J. F., Gaertner, S. E., Kawakami, K., & Hodson, G. (2002). Why can't we just get along? Interpersonal biases and interracial distrust. *Cultural Diversity and Ethnic Minority Psychology*, *8*, 88.

Durkheim, E. (1893). *The division of labor in society*. New York, NY: Macmillan.

Durkheim, E. (1915). *The elementary forms of the religious life*. London, England: G. Allen & Unwin.

Ellemers, N., Kortekaas, P., & Ouwerkerk, J. W. (1999). Self–categorisation, commitment to the group and group self–esteem as related but distinct aspects of social identity. *European Journal of Social Psychology*, *29*, 371–389.

Fazio, R. H. (2001). On the automatic activation of associated evaluations: An overview. *Cognition & Emotion*, *15*, 115–141.

Feeney, J. A., & Noller, P. (1990). Attachment style as a predictor of adult romantic relationships. *Journal of Personality and Social Psychology*, *58*, 281–291.

Festinger, L. (1962). *A theory of cognitive dissonance* (*Vol. 2*). Stanford, CA: Stanford University Press.

Fischer, R., Callander, R., Reddish, P., & Bulbulia, J. (2013). How do rituals affect cooperation? *Human Nature*, *24*, 115–125.

Freeman, L. C., & Webster, C. M. (1994). Interpersonal proximity in social and cognitive space. *Social Cognition*, *12*, 223–247.

Gallup, A. C., Hale, J. J., Sumpter, D. J., Garnier, S., Kacelnik, A., Krebs, J. R., & Couzin, I. D. (2012). Visual attention and the acquisition of information in human crowds. *Proceedings of the National Academy of Sciences*, *109*, 7245–7250.

Gray, K., Rand, D. G., Ert, E., Lewis, K., Hershman, S., & Norton, M. I. (2014). The emergence of "us and them" in 80 lines of code modeling group genesis in homogeneous populations. *Psychological Science*, *25*, 982–990.

Greenwald, A. G., McGhee, D. E., & Schwartz, J. L. (1998). Measuring individual differences in implicit cognition: The implicit association test. *Journal of Personality and Social Psychology*, *74*, 1464–1480.

Greenwald, A. G., Nosek, B. A., & Banaji, M. R. (2003). Understanding and using the implicit association test: I. An improved scoring algorithm. *Journal of Personality and Social Psychology*, *85*, 197–216.

Hamilton, W. D., & Axelrod, R. (1981). The evolution of cooperation. *Science*, *211*, 1390–1396.

Halberstadt, J., Jackson, J. C., Bilkey, D., Jong, J., Whitehouse, H., McNaughton, C., & Zollmann, S. (2016). Incipient social groups: An analysis via in-vivo behavioral tracking. *PloS One*, *11*, e0149880.

Hayduk, L. A. (1983). Personal space: Where we now stand. *Psychological Bulletin*, *94*, 293–335.

Holmes, J. D. (2009). Transparency of self-report racial attitude scales. *Basic and Applied Social Psychology*, *31*, 95–101.

Jackson, J. C., Rand, D., Lewis, K., Norton, M. I., & Gray, K. (2017). Agent-based modeling: A guide for social psychologists. *Social Psychology and Personality Science*. doi:10.1177/1948550617691100

Knobloch, L. K., Solomon, D. H., & Cruz, M. G. (2001). The role of relationship development and attachment in the experience of romantic jealousy. *Personal Relationships*, *8*, 205–224.

Latané, B. (1981). The psychology of social impact. *American Psychologist*, *36*, 343–356.

Levitt, S. D., & List, J. A. (2007). What do laboratory experiments measuring social preferences reveal about the real world? *The Journal of Economic Perspectives*, *21*, 153–174.

Lewis, J. P. (1995). Fast template matching. *Vision Interface*, *95*, 15–19.

Macy, M., & Flache, A. (2009). Social dynamics from the bottom up: Agent-based models of social interaction. In P. Headstrom & P. Bearman (Eds.), *The Oxford handbook of analytical sociology* (pp. 245–268). Oxford, England: Oxford University Press.

Mayseless, O., & Scharf, M. (2007). Adolescents' attachment representations and their capacity for intimacy in close relationships. *Journal of Research on Adolescence*, *17*, 23–50.

McClure, M. J., & Lydon, J. E. (2014). Anxiety doesn't become you: How attachment anxiety compromises relational opportunities. *Journal of Personality and Social Psychology*, *106*, 89.

Mehl, M. R., & Holleran, S. E. (2007). An empirical analysis of the obtrusiveness of and participants' compliance with the electronically activated recorder (EAR). *European Journal of Psychological Assessment*, *23*, 248–257.

Mehl, M. R., Pennebaker, J. W., Crow, D. M., Dabbs, J., & Price, J. H. (2001). The electronically activated recorder (EAR): A device for sampling naturalistic daily activities and conversations. *Behavior Research Methods, Instruments, & Computers*, *33*, 517–523.

Milgram, A. (1974). *Obedience to authority*. New York, NY: Harper & Row.

Norenzayan, A., Shariff, A. F., Gervais, W. M., Willard, A. K., McNamara, R. A., Slingerland, E., & Henrich, J. (2016). The cultural evolution of prosocial religions. *Behavioral and Brain Sciences*, *39*, e1.

Onnela, J. P., Waber, B. N., Pentland, A., Schnorf, S., & Lazer, D. (2014). Using sociometers to quantify social interaction patterns. *Scientific Reports*, *4*, *1–8*.

Olguín, D. O., Paradiso, J. A., & Pentland, A. (2006, October). Wearable communicator badge: Designing a new platform for revealing organizational dynamics. In *Proceedings of the 10th International Symposium on Wearable Computers* (pp. 4–6). Montreux, Switzerland: IEEE.

Olguín, D. O., & Pentland, A. S. (2006, October). *Human activity recognition: Accuracy across common locations for wearable sensors*. In Proceedings of 2006 10th IEEE International Symposium on Wearable Computers, Montreux, Switzerland, pp. 11–14.

Olguín, D. O., & Pentland, A. S. (2008). Social sensors for automatic data collection. *AMCIS 2008 Proceedings*, *171*, 1–10.

Payne, B. K., Cheng, C. M., Govorun, O., & Stewart, B. D. (2005). An inkblot for attitudes: Affect misattribution as implicit measurement. *Journal of Personality and Social Psychology*, *89*, 277–293.

Parkinson, B. W., & Spilker, J. J., Jr. (Eds.). (1996). *Progress in Astronautics and Aeronautics: Global positioning system: Theory and applications* (*Vol. 2*). Washington, DC: AIAA.

Pentland, A. (2005). Socially aware, computation and communication. *Computer*, *38*, 33–40.

Porikli, F. (2005, June). Integral histogram: A fast way to extract histograms in Cartesian spaces. In *2005 IEEE Computer Society Conference on Computer Vision and Pattern Recognition (CVPR'05)* (*Vol. 1*, pp. 829–836). Providence, RI: IEEE.

Rosen, M. A., Dietz, A. S., Yang, T., Priebe, C. E., & Pronovost, P. J. (2014). An integrative framework for sensor-based measurement of teamwork in healthcare. *Journal of the American Medical Informatics Association*, *22*, 11–18.

Sardon, E., Rius, A., & Zarraoa, N. (1994). Estimation of the transmitter and receiver differential biases and the ionospheric total electron content from Global Positioning System observations. *Radio Science*, *29*, 577–586.

Schelling, T. C. (1971). *Micromotives and macrobehavior*. New York, NY: WW Norton & Company.

Shaver, P. R., Schachner, D. A., & Mikulincer, M. (2005). Attachment style, excessive reassurance seeking, relationship processes, and depression. *Personality and Social Psychology Bulletin*, *31*, 343–359.

Shteynberg, G. (2010). A silent emergence of culture: The social tuning effect. *Journal of Personality and Social Psychology*, *99*, 683.

Smith, E. R., & Conrey, F. R. (2007). Agent-based modeling: A new approach for theory building in social psychology. *Personality and Social Psychology Review*, *11*, 87–104.

Tajfel, H. E. (1978). *Differentiation between social groups: Studies in the social psychology of intergroup relations*. London, England: Academic Press.

Tajfel, H. (1981). *Human groups and social categories: Studies in social psychology*. Cambridge, MA: CUPArchive.

Tajfel, H. (1982). Social psychology of intergroup relations. *Annual Review of Psychology*, *33*, 1–39.

Tajfel, H. (2010). *Social identity and intergroup relations*. Cambridge, MA: Cambridge University Press.

Tsoar, A., Nathan, R., Bartan, Y., Vyssotski, A., Dell'Omo, G., & Ulanovsky, N. (2011). Large-scale navigational map in a mammal. *Proceedings of the National Academy of Sciences*, *108*, E718–E724.

Turner, J. C. (1982). Towards a cognitive redefinition of the social group. In H. Tajfel (Ed.), *Social identity and intergroup relations* (pp. 15–40). Cambridge, England: Cambridge University Press.

Wahlberg, M. (2002). The acoustic behaviour of diving sperm whales observed with a hydrophone array. *Journal of Experimental Marine Biology and Ecology*, *281*, 53–62.

Want, R., Hopper, A., Falcao, V., & Gibbons, J. (1992). The active badge location system. *ACM Transactions on Information Systems (TOIS)*, *10*, 91–102.

Whitehouse, H. (2012). 10 Ritual, cognition, and evolution. In R. Sun (Ed.), *Grounding social sciences in cognitive sciences* (pp. 265–284). Cambridge, MA: MIT Press.

Whitehouse, H., & Lanman, J. A. (2014). The ties that bind us. *Current Anthropology*, *55*, 674–695.

Wiltermuth, S. S., & Heath, C. (2009). Synchrony and cooperation. *Psychological Science*, *20*, 1–5.

Winking, J., & Mizer, N. (2013). Natural-field dictator game shows no altruistic giving. *Evolution and Human Behavior*, *34*, 288–293.

Zhan, B., Monekosso, D. N., Remagnino, P., Velastin, S. A., & Xu, L. Q. (2008). Crowd analysis: A survey. *Machine Vision and Applications*, *19*, 345–357.

Author Biographies

Joshua Conrad Jackson is a doctoral student at University of North Carolina at Chapel Hill. He studies culture, religion, and morality, among other topics.

David Bilkey is a professor of psychology and the director of the Brain Health Research Centre at the University of Otago. His primary research investigates the biological basis of memory and how memory-associated brain regions such as the hippocampus are altered in schizophrenia.

Jonathan Jong is a research fellow at Coventry University and Research Associate at the University of Oxford. His research focuses on the cognitive and cultural foundations of religion and morality.

Maya Rossignac-Milon is a doctoral student at Columbia University. Her research explores the role of shared reality and epistemic motives in fostering interpersonal closeness both between strangers and romantic relationship partners.

Jamin Halberstadt is a professor of psychology at the University of Otago in Dunedin, New Zealand. His interests include, among others, the interface of emotion and cognition, social categorization, and issues in religious cognition. He is a senior editor at the journal *Psychological Science*.

Handling Editor: Lorne Campbell and Simine Vazire