PHILOSOPHICAL TRANSACTIONS B

royalsocietypublishing.org/journal/rstb

Opinion piece

Cite this article: Gratch J. 2023 The promise and peril of interactive embodied agents for studying non-verbal communication: a machine learning perspective. *Phil. Trans. R. Soc. B* **378**: 20210475.
https://doi.org/10.1098/rstb.2021.0475

Received: 30 June 2022
Accepted: 27 January 2023

One contribution of 15 to a discussion meeting issue 'Face2face: advancing the science of social interaction'.

Subject Areas:
cognition, computational biology, behaviour

Keywords:
experimental methodology, causality, computational models

Author for correspondence:
Jonathan Gratch
e-mail: gratch@ict.usc.edu

THE ROYAL SOCIETY
PUBLISHING

The promise and peril of interactive embodied agents for studying non-verbal communication: a machine learning perspective

Jonathan Gratch

Department of Computer Science, University of Southern California, Los Angeles, CA 90292, USA

JG, 0000-0002-5959-809X

In face-to-face interactions, parties rapidly react and adapt to each other's words, movements and expressions. Any science of face-to-face interaction must develop approaches to hypothesize and rigorously test mechanisms that explain such interdependent behaviour. Yet conventional experimental designs often sacrifice interactivity to establish experimental control. Interactive virtual and robotic agents have been offered as a way to study true interactivity while enforcing a measure of experimental control by allowing participants to interact with realistic but carefully controlled partners. But as researchers increasingly turn to machine learning to add realism to such agents, they may unintentionally distort the very interactivity they seek to illuminate, particularly when investigating the role of non-verbal signals such as emotion or active-listening behaviours. Here I discuss some of the methodological challenges that may arise when machine learning is used to model the behaviour of interaction partners. By articulating and explicitly considering these commitments, researchers can transform 'unintentional distortions' into valuable methodological tools that yield new insights and better contextualize existing experimental findings that rely on learning technology.

This article is part of a discussion meeting issue 'Face2face: advancing the science of social interaction'.

1. Introduction

Face-to-face social interaction is the most ubiquitous and rewarding of human activities, yet in many ways, the science of face-to-face social interaction is still in its infancy. Science, at its core, 'should detail the cogs and wheels of the causal process through which the outcome to be explained was brought about' [1, p. 50]. While detailing these mechanisms in the physical sciences can be enormously challenging, at least physical processes do not have minds that learn new mechanisms in response to external stimuli. By contrast, parties in a face-to-face interaction form and change beliefs about their partner and the social context, adjust their behaviour and thereby jointly create the outcome to be explained. This fundamentally *interdependent* nature of social actors creates problems for traditional controlled experiments used to establish causal mechanisms in the physical sciences.

This article reviews some of the methodological limitations that arise from the most common approaches used to establish the mechanisms underlying face-to-face interaction, then highlights the emerging paradigm of using embodied interactive agents—such as virtual humans [2–4] or social robots [5]—as a potential way to avoid these limitations. Developers of embodied agents increasingly rely on machine learning methods to create high-fidelity models of non-verbal behaviour [6] and verbal communication [7] by analysing large datasets of face-to-face interaction. Ironically, the use of machine learning can inadvertently reintroduce the very limitations these agents were intended to avoid. Essentially, by their choice of learning method, the selection of

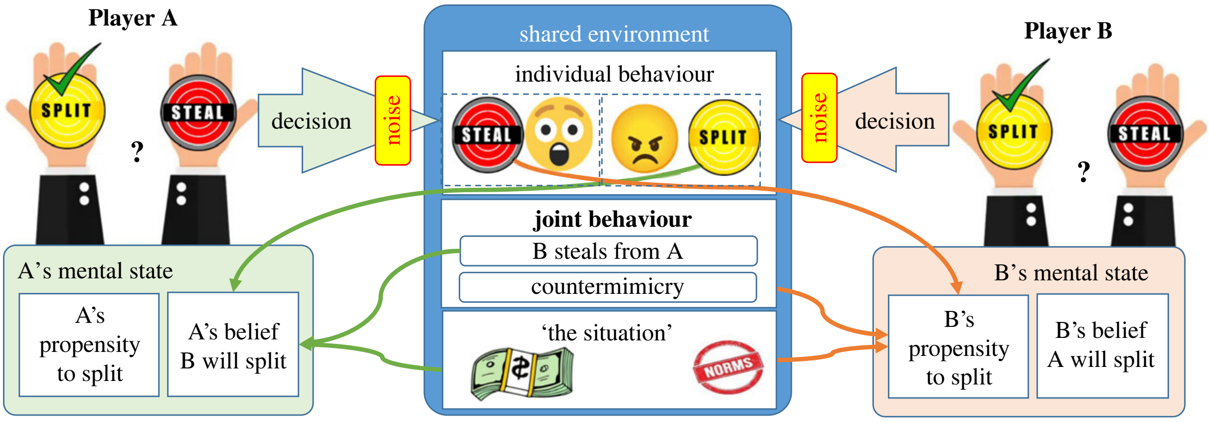


Figure 1. The *noisy* iterated Prisoner's Dilemma game. Each round, players simultaneously decide if they will try to split or steal a pot of money but their choice may be imperfectly executed (e.g. Player A intends to split but steal is performed). Mechanisms for modelling the interaction between players can be distinguished by their inputs (e.g. do decisions to split only depend on features of the environment, do they depend only on the partners actions, do they depend on joint behaviours?) and whether they posit internal state such as beliefs about the partner or beliefs about the partner's beliefs. (Online version in colour.)

features, and the choice of training protocol, agent designers may unintentionally distort the mechanisms they seek to model. Further, owing to the opaque nature of some machine learning methods, these distortions can be difficult to recognize. When done with foresight, however, the pallet of machine learning design choices can serve as fertile ground for hypothesizing and testing alternative interaction mechanisms (e.g. [8]). Here, I highlight how these issues arise in the study of the moment-to-moment emotional expressions and other non-verbal signals that are the hallmark of face-to-face interaction.

The fundamental challenge in studying face-to-face interaction is that parties are interdependent. This observation goes back at least as far as gestalt psychologist Kurt Lewin, who argued that group interactions are best viewed as a *dynamic whole* such that the change in the state of any member changes the state of any other member [9]. Interdependence arises even in seemingly one-sided interactions, as when one person tells a story to a silent listener. As Janet Bavelas *et al.* illustrated when testing her collaborative theory of storytelling, the gaze and facial expressions of active listeners serve to shape and co-construct the narrator's story [10]. Studies with embodied agents have replicated this effect [11], highlighting the promise of this technology.

In seeking to explain how a narrator and silent listener can co-construct a story, or how two negotiators can converge on a mutually beneficial deal, a science of face-to-face interaction must hypothesize mechanisms that underlie such interdependent behaviour and develop valid empirical and analytic methods to contrast alternative hypotheses. Some research emphasizes mechanisms that reside within the mind of each participant. For example, social neuroscience seeks to find the neural circuits that underlie social encounters [12] and cognitive psychology posits abstract algorithms that 'run' in the brain [13,14]. Other work has emphasized mechanisms in the environment. For example, social and organizational psychology investigates how roles, norms and rules shape how interactions unfold [15], and embodied theories of cognition highlight how organisms off-load cognitive work onto the environment [16,17]. Some goes so far as to argue that social interaction must be studied as a thing unto itself and cannot be reduced to individual cognitive or environmental mechanisms [18,19].

To illustrate these possible perspectives, consider figure 1, wherein two players, A and B, are engaged face-to-face in a multi-round *noisy* Prisoner's Dilemma [20]. The classic Prisoner's Dilemma creates a conflict between self-interest (stealing all or most of a pot of money on the table) and collective interest (fairly dividing this money with one's partner). Players privately record their intention to split or steal (though the choice is often labelled in less obvious terms), and the reward in each round is determined by the players' joint decision. The noisy version of the game allows for discrepancies between the intended and actual outcome owing to unintended errors (e.g. Player A may choose to split the money but, with some probability, a steal action is performed by the game). This creates ambiguity as to a party's true intent, which might be resolved through communication (verbal or non-verbal) and observation of a player's pattern of behaviour over time. Even without noise, ambiguity can arise as to whether Player A stole out of a greedy disposition or as retribution for Player B's past behaviour, and human players often fail to account for how their prior actions can determine their partner's response when resolving such ambiguity [21].

A science of face-to-face interaction should be able to hypothesize mechanisms that explain how players coordinate their actions to achieve win–win solutions (or successfully exploit their partner) and discern if these mechanisms demand complex cognitive machinery like Theory of Mind [22], or if this complexity can be off-loaded to the environment or features of the interaction itself. Indeed, it is possible to model human behaviour in the Prisoner's Dilemma by positing simple stimulus–response rules that are triggered by observable features in the environment.

For example, Axelrod's famous Tit-for-Tat mechanism simply matches the previously observed action of the partner (ignoring the player's own behaviour). If Player A steals, Player B steals in the next round. In one study of a large corpus of face-to-face Prisoner Dilemma games, Tit-for-Tat predicted player decisions with 70% accuracy, whereas a machine learning model that incorporated a player's history of actions and facial expressions only improved this to 74% [23]. Other work argues that interactional features (which cannot be inferred from looking at one individual alone) are crucial. For example, people have been argued to automatically mimic each other's behaviour [24] or 'catch' each other's physiological state [25], and these patterns have been shown useful for predicting interpersonal outcomes [26,27].

Such simple models are compelling but also ignore decades of research highlighting how behaviour often involves rich mental representations, such as ascribing mental states to other people to explain and predict their actions, or what is called Theory of Mind [22]. Indeed, the noisy version of the Prisoner's Dilemma was developed to illustrate how causal ambiguity (a defining characteristic of everyday interactions) undermines the predictive value of simple mechanisms like Tit-for-Tat [20]. Rather, theories of how people interact in such tasks typically assume players are fully intentional agents with Theory of Mind. For example, Kelley [21] posited that each player has an *a priori* propensity to share or steal—sometimes called their social value orientation [28]—and simultaneously tries to infer their partner's propensity when making a decision. Thus, Player B might see Player A's facial expression of surprise, not as arising from automatic mimicry or contagion, but as a communicative assertion that Player A's act of stealing was unintended [29]. And more generally, interactions might unfold and shift as players separately attend to their own and their partner's history of behaviour, transformed by their *a priori* biases and beliefs about their partner's biases.

Given this diversity of possible mechanisms, how can a science of face-to-face interaction progress? In this paper, I advocate a 'learned-partner' approach where hypothetical mechanisms are learned from face-to-face human interactions and then evaluated using experiments that incorporate these learned mechanisms into interactive social agents. Before introducing this approach, I first review why establishing causality is difficult when parties are interdependent and highlight the limitations of traditional experimental methods. While a promising approach to address the challenges of traditional methods, I will argue that these 'learned-partners' may incorporate design decisions that unintentionally distort the very interactivity they seek to illuminate. I will discuss an ontology of the (often implicit) theoretical commitments that agent-based models incorporate and their implications for a science of interaction. Articulating these commitments can help realize the potential of learned-partners by guiding computer science research and better contextualizing existing experimental findings that rely on learning technology. While the fundamental tension between interactive and experimental control is present in any domain of social interaction, I approach these issues from the perspective of behavioural game theory and the role of non-verbal communication in shaping behaviour in the type of economic games typically studied in that literature, and the issues and solutions may play out differently in other social contexts.

2. Establishing causality in interdependent actors

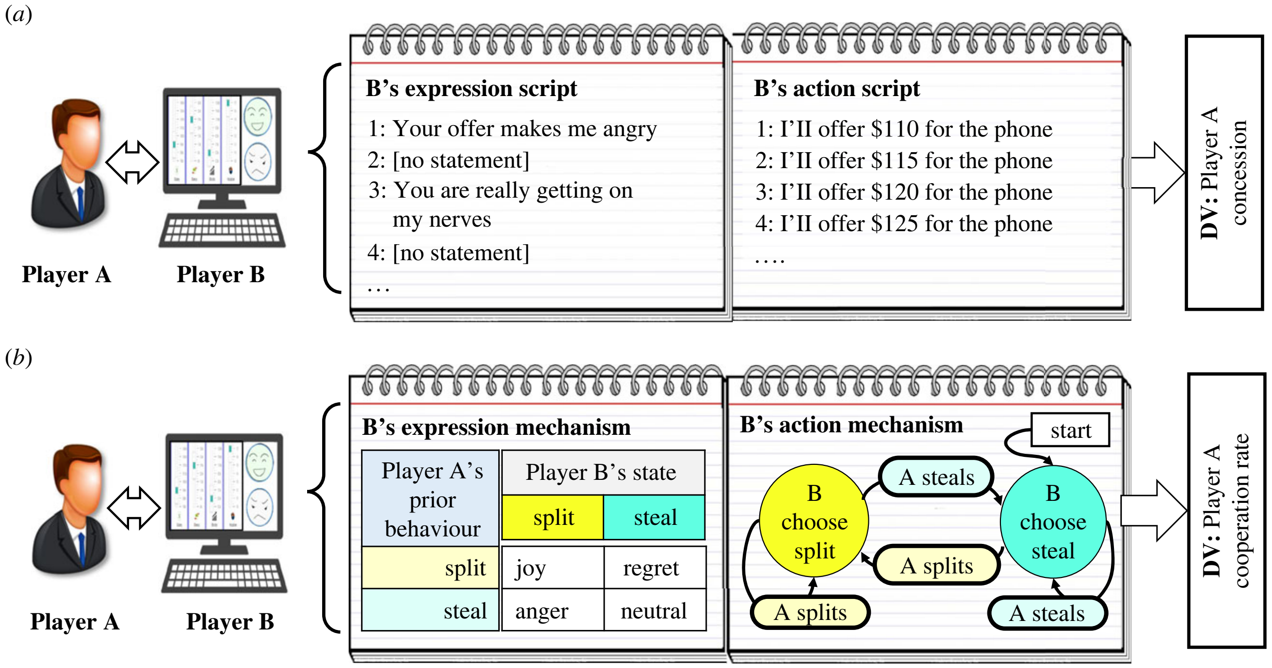
According to the traditional scientific method, several criteria must be satisfied to establish a causal relationship [30]. First, research must establish an *association* (e.g. Player B's propensity to steal is correlated with Player A's propensity to steal). Second, one must show *temporal precedence* (e.g. Player B only stole after Player A stole first). Third, one must show the association is *non-spurious.* An association is spurious if it is due to changes in a third factor (e.g. both players' propensity to steal is shaped by the time of day, see [31]). Finally, most scientists argue that a causal explanation is inadequate unless evidence is provided for the specific *mechanism* that creates a connection between variation in an experimental manipulation and variation in the dependent measure (e.g. the 'cogs and wheels').

Conventional 'gold standard' experimental methods (such as randomized control trials or A/B testing) contain several features that help establish these criteria. First, experiments include at least two comparison groups that differ according to some theoretically relevant factor (this difference is called the independent variable). Second, some outcome measure is defined (called the dependent variable). Third, participants are randomly assigned to groups to rule out spurious associations between the independent and dependent variable. Many designs include additional measures or manipulations to explicitly test mechanisms, such as showing the relationship between independent and dependent variables vanishes when a hypothesized causal mechanism is controlled for [32] or manipulated via experimental design [33]. As an example, an experiment could examine if seeing one's partner's facial expressions (the independent variable) impacts propensity to steal (the dependent variable). Participants could play a 10-round Prisoner's Dilemma over Zoom and randomly assigned to playing with the video on or off, and the number of times either player steals could be the dependent measure. A hypothesized mechanism might be that expressions shape the perceived trustworthiness of the partner, which could be measured via self-report after the game and then examined via mediation analysis [34].

When it comes to the study of face-to-face interaction, however, these features can be difficult to enforce without distorting or disrupting the very mechanisms one wishes to study. To establish the causal mechanisms underlying face-to-face interaction, social psychologists overwhelmingly rely on a small set of experimental methods. In *group experiments*, two or more participants engage in a social interaction where the collective is treated as the unit of analysis: i.e. independent variables impact the entire group and dependent variables are defined in terms of group behaviours or outcomes. In *non-contingent experiments*, the individual is treated as the unit of analysis: a participant engages with a social stimulus that does not respond to (i.e. is not contingent upon) the participant's behaviour. This could simply involve asking a participant's impressions from watching a pre-recorded video (so-called 'spectator experiments', see [35]), or a participant interacting with a confederate that produces scripted behaviours. *Contingent experiments* (sometimes called 'second person' experiments, see [12]) allow individual participants to socially engage with some contingent stimuli (e.g. play Prisoner's Dilemma with a Tit-for-Tat strategy), where designs independently manipulate alternative contingent mechanisms. I briefly review limitations of each.

(a) Group experiments

In group experiments, multiple individuals interact naturally but aspects of the situation are manipulated to illuminate the mechanisms underlying interdependent behaviour. For example, by assigning complementary versus competitive goals, one can show that negotiators become more synchronized in posture and gesture [36], and infants and mothers



B's expression script
1: Your offer makes me angry
2: [no statement]
3: You are really getting on my nerves
4: [no statement]
…

B's action script
1: I'Il offer $110 for the phone
2: I'Il offer $115 for the phone
3: I'Il offer $120 for the phone
4: I'Il offer $125 for the phone
….

B's expression mechanism

Player A's prior behaviour	Player B's state	
	split	steal
split	joy	regret
steal	anger	neutral

Figure 2. Non-contingent experiments (*a*) follow a fixed script to ensure Player B's behaviour (offers and emotional expressions) is guaranteed independent of Player A's prior behaviour. By contrast, contingent experiments (*b*) incorporate mechanisms that respond conditionally to Player A's prior (Player B's expressions and choices to cooperate are conditioned on Player A). Each approach can treat one or more of these scripts/mechanisms as independent variables. DV, dependent variable. (Online version in colour.)

exhibit more neural synchrony when they can engage in mutual gaze, compared with a separate group of dyads that can not [37]. It is tempting to conclude these group changes tell us something of the nature of mechanisms within the mind of each participant. For example, perhaps negotiators mimic their partner when they are in a cooperative context but fail to mimic when they find themselves in competition [38,39]. Yet all one can firmly conclude is that the experimental manipulation is causing a change in the group. For example, friends will show more neural synchrony than strangers when watching a debate alone in their own home, not because they respond to each other but because they are interpreting the same environmental events with similar goals [40]. Similarly, the fact that negotiators with complementary goals become synchronized could simply reflect their reactions to a shared task.

The problem with inferring individual mechanisms from group behaviour is that features of the interaction, like synchrony, simply provide evidence of an association, but to show causality it is necessary (though not sufficient) to establish temporal precedence. Thus, there is growing interest in the use of dynamic systems methods to show temporal precedence from dyadic data. For example, actor–partner methods [41,42] or lag analysis [43] can show, for example, that if Player A in the Prisoner's Dilemma smiles, Player B tends to also smile, with a small delay. This temporal precedence suggests Player A is the leader and Player B is adapting their own behaviour in response. Using such methods, Mendes and colleagues examined mixed-race dyads and argued that African-Americans 'catch' the anxiety of their European–American conversation partners, but not *vice versa*, and thus concluded minorities are more vigilant of, and more likely to adjust to their partners. While compelling, the fundamental interdependence between partners makes it hard to rule out spurious or misleading temporal dependencies, leading others to the comfort of more traditional designs when trying to make inferences about the individual.

(b) Non-contingent experiments

Non-contingent experiments allow firm inferences about the individual by sacrificing the interdependence of natural face-to-face interactions. Specifically, if we wish to claim that some measure of Player A in the Prisoner's Dilemma (our dependent variable) is *caused* by some behaviour of Player B (our independent variable), we must ensure that the behaviour of Player B is not, in turn, dependent on prior behaviours of Player A. Using this approach, Van Kleef *et al*. [44] sought to establish that expressions of anger by a negotiator *cause* concessions in the negotiator's partner. Each participant (Player A) engaged in a negotiation with what they believed to be another participant (Player B) over the Internet (figure 2*a*). Player A took turns sending and receiving offers and text messages with Player B. In fact, Player B was a scripted computer program that produced angry (or happy) statements on pre-programmed rounds (e.g. 'Your offer makes me angry'), and made concessions according to a pre-programmed schedule (e.g. offer $115 on round two and $120 on round three, etc.). The experiment showed that participants concede more to an angry opponent. Van Kleef *et al.* also provided evidence for the mechanism. They hypothesized Player B's anger acts as a signal that alters the participant's belief in Player B's willingness to concede. This was supported via statistical mediation on subjective reports about Player B's limits.

Unfortunately, this increase in control over the individual can destroy the very interactivity one hopes to study. For example, one of the hallmarks of good negotiators is that they can adjust to each other and co-construct mutually beneficial solutions [45]. In a salary negotiation, both sides might think they bring opposing interests, but through building trust and reciprocally revealing information about their goals, they might discover that one side would welcome a lower salary in exchange for a more flexible work schedule. Study of such emergent processes is simply not possible if one side is ignoring the other.

Perhaps a more fundamental concern for a science of interaction, non-contingent experiments lead people to believe they are engaged in social interaction when, in fact, their attempts to interact are ignored. Indeed, non-contingency may unintentionally introduce some new factor to the experiment—e.g. the lack of contingency could signal their partner is higher in power [39] or unnatural [46], thus reducing ecological validity. Indeed, studies in social neuroscience suggest that people are highly sensitive to the contingency of their partner's behaviour and engage different neural mechanisms in the presence of contingent partners [12,47]. Related to this, non-contingent partners may also reduce some aspects of experimental control. Specifically, non-contingent design may fail to control *subjective* variables essential to understand *interdependent* interactions as our expectations of a partner may be formed relative to our own prior behaviour. For example, imagine Player A has made an unfairly high offer. Player B's resulting anger seems understandable given Player A's norm violation, and we might expect Player A to repair this norm violation by making a less ambitious subsequent offer. Imagine instead, Player A has made a very generous offer and Player B responds with anger. Now it is Player B's behaviour that violates social norms. What this illustrates is the *meaning* of Player B's anger may be contingent on Player A's immediately preceding action [48]. But scripted agents seek to control the act (e.g. expressing anger), but not the meaning this act has for the observer.

(c) Contingent experiments

An alternative to non-contingent experiments, less common in social psychology but popular in computer science and game theory (e.g. [8,49]), is to introduce and independently manipulate the *contingency* of partner behaviour. For example, imagine a virtual human has been programmed to follow the gaze of a participant while they are looking at different objects on a table [50]. By independently manipulating alternative mechanisms (e.g. follow gaze versus random gaze), these designs seek to make valid causal inferences about the consequences of interacting with specific mechanisms (e.g. joint attention is needed to activate certain reward circuits in the brain, see [51]).

Contingent designs have greater ecological validity for the study of social interaction as they better approach the natural interactivity of face-to-face interaction and they better control the *meaning* of the partner's action (e.g. a contingent negotiation agent might show anger only if Player A makes an unfair offer). On the negative side, they suffer an opposite issue that can reduce experimental control over the partner's actions. To see this, consider an experiment where participants play the Prisoner's Dilemma with a partner that mimics their facial expressions—smiles when they smile; frowns when they frown—versus participants that play with a partner that engages in counter-mimicry—frowns when they smile; smiles when they frown [52]. Based on their propensity to smile, participants are essentially self-selecting the expressions of their contingent partner. While the design still allows valid conclusions across experimental conditions (e.g. people smile more when interacting with a partner that mimics their smiles), contingent designs can amplify the variance arising from participants' individual differences, requiring larger sample sizes and additional experimental tests, such as statistically controlling for these differences, to rule out spurious mechanisms. Practically speaking, contingent designs may be more difficult for human confederates to execute, allowing experimenter effects [53] to creep into designs (wherein they unwittingly deviate from the intended script). As a result, contingent designs typically involve simple mechanisms (e.g. Tit-for-Tat) that, while improving on non-contingent designs, may lack the nuance of actual human behaviour and violate expectations that participants acquire from years of socialization.

Figure 2*b* illustrates an example of a contingent design examining the interactional function of facial expressions [48]. Participants were told they were playing a social task (the iterated Prisoner's Dilemma) with another participant whose expressions were reflected on a graphical avatar. In truth, a computer controlled the avatar's expressions and actions, but rather than following a deterministic script, the partner's emotional expressions and task decisions were contingent on the participant's actions. The automaton on the right of figure 2*b* dictates the computer-controlled partner's decisions follow a Tit-for-Tat strategy: the computer-controlled partner chooses 'split' as long as the participant chooses 'split', but switches to stealing if the participant chooses to steal, and *vice versa*. The expressions are similarly contingent: if the computer-controlled partner chooses 'split' and the participant 'split' (i.e. mutual cooperation), the computer will express joy; if the partner chooses to steal, the computer will express anger.

Using such mechanisms, one can perform experiments comparing how participants play against alternative contingent mechanisms. For example, participants will cooperate more with the expression policy that expresses regret after successfully stealing from the participant (as shown in figure 2*b*) compared with an expression policy that expresses joy [48]. Perhaps most importantly, a focus on mechanisms enables the study of interdependent behaviours and outcomes found in group experiments, while maintaining some measure of experimental control over the individual. For example, in group experiments, Thompson [45] finds that participants discover win–win deals by exchanging information and concurrently adapting their offers to better account for their partner's interests, and this effect can be replicated with contingent agents that adapt their offers to the human participant [54]. Further, it is possible to create more nuanced mechanisms to explore, for example, how the discovery of win–win solutions is impacted by different communication strategies—such as if a person is willing to freely share their most important priorities, or if this sharing should be contingent on sharing by the partner [54]—and how it is shaped by hypothesized cognitive biases such as the anchoring or the fixed-pie bias [55].

(d) Discussion

Group and non-contingent experiments are important tools for uncovering the cogs and wheels of face-to-face interactions, but each is problematic for those that seek to uncover the cognitive and neural mechanisms that shape interdependent behaviour. Dyadic experiments preserve the interdependent nature of social interactions but only allow valid conclusions about behaviours of the group. Non-contingent experiments allow strong

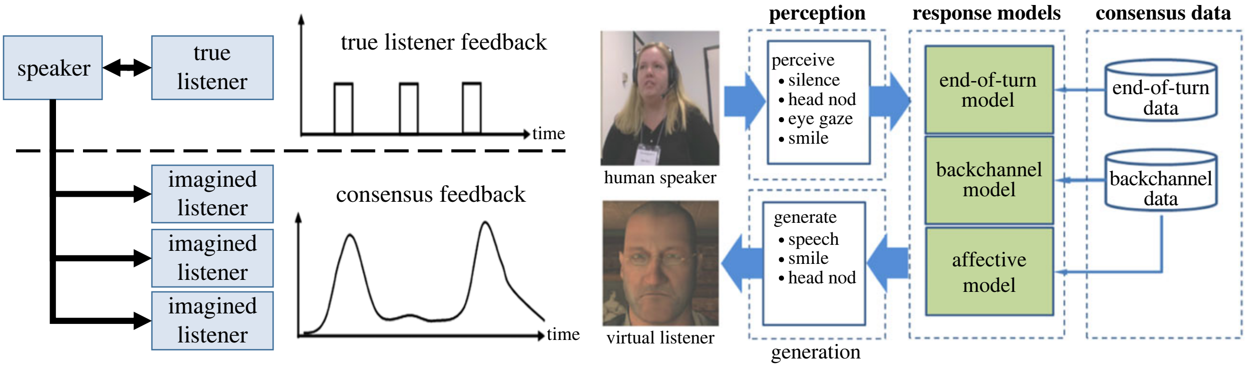


Figure 3. An active-listening agent [74] incorporates models of a listener's turn-taking, backchannels and facial expressions. The model was trained by having multiple listeners imagine they were in a conversation with a pre-recorded video, thereby examining individual variability in listener behaviour. (Online version in colour.)

experimental control over the individual, but in doing so, they break the interdependent nature of face-to-face interactions. Thus, they may undermine participants' sense that they are engaged in social interaction, lower ecological validity and suffer the criticisms of 'spectator experiments' [12,35]. Contingent experiments allow the study of interdependent behaviour, but together with non-contingent experiments, often rely on simple mechanisms that strip away much of the complexity of actual human behaviour.

3. 'Learned-partner' experiments

One approach to hypothesizing complex interaction mechanisms is to learn them from examples of people engaged in natural face-to-face interaction. Recent advances in machine learning are revolutionizing the development of embodied (and disembodied) interactive agents. This is evident with the rapid advancement of conversational assistants like Alexa, but the fields of Affective Computing and Multimodal Interaction are also making dramatic progress in recognizing and predicting non-verbal interactional behaviours. Barquero and colleagues [6] recently surveyed a large number of machine learning approaches that analyse non-verbal behaviour in face-to-face interactions. Such approaches develop models, for example, of when and how listeners provide non-verbal feedback to speakers [56,57], how non-verbal cues shape turn-taking behaviour [58] and what pattern of behaviours predict subsequent engagement or disengagement [59].

For the purpose of this article, these models can be seen both as candidate mechanisms (i.e. the cogs and wheels) of human interaction—as they were inferred through an analysis of natural interactions—and as mechanisms that can be incorporated into embodied agents to enhance the realism and ecological validity of contingent experiments. This does not come without challenges, as I will briefly revisit in the conclusion. For example, models learned purely from behaviour are difficult to interpret and may discover mechanisms that lack biological plausibility unless guided by theory (e.g. using theoretically posited features or representations) or 'nudged' to align with internal biological mechanisms [60]. Setting these complications aside for now, I first illustrate an example machine learning approach to modelling non-verbal communication before taking a more critical look at the assumptions and potential pitfalls of uncritically using these models to study human interdependent behaviour, including for example, if they misrepresent the necessity of Theory of Mind in understand the function of such non-verbal signals.

(a) An example learned-partner: active-listening agents

In Barquero *et al.*'s [6] review of non-verbal prediction models, by far the most common area of progress is simulating active-listening behaviours such as 'backchannel' feedback and smooth turn-taking found in conversations. Backchannels are verbal (e.g. *yes*), vocal (e.g. *uh-huh*) and/or gestural (head movements, eyebrows movements, smiles) produced by listeners, and associated with 'backchannel-invitation cues' produced by speakers, such as gaze patterns, pauses and changes in vocal prosody [61]. Backchannels convey the listener is actively engaged and how they are evaluating the speaker's words. Speakers adjust to this feedback, thus speaker and listener are interdependent partners that co-construct the conversation [10,62]. Such behaviour has been studied under many names, including 'rapport' [63], 'social resonance' [64,65], 'interpersonal adaptation' [66], 'entrainment' [67], 'interactional synchrony' [68], 'social glue' [69], 'immediacy behaviours' [70] and 'positivity resonance' [71]. From an interaction standpoint, these behaviours help to promote more effective and persuasive communication (e.g. [72,73]) and increase subjective feelings of rapport [63].

Figure 3 illustrates one approach to modelling active listeners from our research group [74,75]. As a starting point, we collected a corpus of quasi-monologues between pairs of participants (one speaker told a story and the listener was instructed not to speak but could freely communicate by non-verbal behaviour such as nods or smiles). From this, we created a set of speaker-only videos with each speaker facing a camera (see 'human speaker' in figure 3). Each of these videos was presented to multiple 'imagined listeners' from a separate group of participants. These were instructed, similar to Schilbach *et al.* [76], to imagine they were in a real social interaction and behave accordingly. Using multiple imagined listeners reveals variability in listener behaviour. For example, certain cues of the speaker might signal a request for feedback (e.g. pause briefly and look at the listener), but differences in the listener's engagement or personality might impact if they respond to this signal. Thus, learning algorithms were trained to predict a measure of consensus across all of the imagined speakers such that the model would predict a high likelihood of feedback where there was strong consensus. See the work

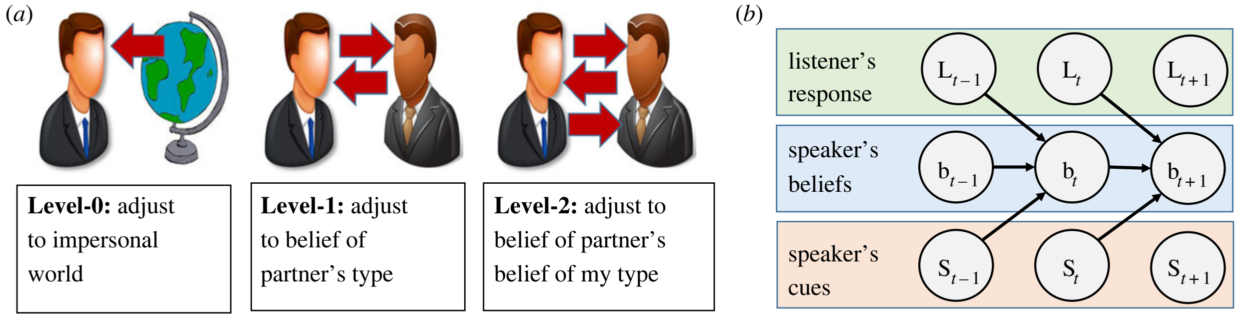

(a)
Level-0: adjust to impersonal world
Level-1: adjust to belief of partner's type
Level-2: adjust to belief of partner's belief of my type
(b)
listener's response
L_{t-1}
L_t
L_{t+1}
speaker's beliefs
b_{t-1}
b_t
b_{t+1}
speaker's cues
S_{t-1}
S_t
S_{t+1}

Figure 4. (*a*) Levels of Theory of Mind in terms of nested beliefs about the interaction partner. (*b*) A model of an active listener that attributes beliefs to the speaker based on cues, when determining a response [88]. (Online version in colour.)

of de Kok & Heylen for a similar approach [77]. Models of backchannelling, turn-taking and smiling were learned independently, though more recent work emphasizes the benefit of training these models concurrently (e.g. [57]).

Several studies have examined how people interact with embodied agents that incorporate active-listening models (both learned and hand-crafted). Our early research with a hand-crafted model found that participants are sensitive to the contingency of listener backchannelling. Participants self-reported more rapport when interacting with a virtual human that generated nods that were contingent on their own speech compared with a *yoked condition* [78] where they saw nods contingent on a previous participant's speech [79]. Behaviourally, they showed more self-disclosure [79,80], more fluent speech [79–81], more mutual gaze [81] and improvement in performance when a contingent agent of opposite sex supervised a maths exam [82,83]. The interactional consequences of learned models have been less rigorously examined—it is more common just to report fit to data or evaluate pre-recorded interactions simulated by the model (e.g. [84])—but a few true interaction studies show similar results. For example, Park and colleagues incorporated a learned backchannel prediction model into an embodied robot and found children gaze significantly more and speak with higher energy to a contingent robot [85].

(b) A critical analysis of active-listening agents

It is tempting to view learned models as black boxes that take features of the interaction as input and, as the field progresses, produce ever-more accurate simulations of real human behaviour. And indeed, just considering active-listening agents, there have been considerable advances in machine learning approaches moving from supervised machine learning techniques which require costly hand annotation [86] to semi-supervised algorithms which can augment human annotation with massive amounts of unlabelled data [84]. But going back to figure 1, it is important to consider the inputs and internal representations these models consider.

Most early and many contemporary active-listening models learn a mapping from the features of the speaker to behaviours of the listener. For example, a simple learned model might suggest that an embodied agent should nod when the speaker pauses for more than 50 ms while gazing at the listener. Such a model treats the speaker as part of the visible environment (it serves as a simple condition–action rule that does not impute any mental states to the speaker) and ignores any internal state of the listener when generating a response. By contrast, the multi-listener corpus collected by Huang and colleagues in figure 3 begins to open up consideration of variance in the listener's behaviour. For example, we examined how listener feedback varies with several personality types including the Big Five (extraversion, agreeableness, openness, conscientiousness, and neuroticism), self-consciousness and self-monitoring. We found that these traits are associated with different patterns of listening behaviour [87]. Thus, by incorporating listener features, it now becomes possible to create different contingent-listeners that approximate different personality types, and one could potentially examine if human speakers recognize or behave differently towards these different mechanisms. It also becomes possible for an automated speaker to use such a model to recognize individual differences in human listeners (by observing patterns in their listening behaviour) and adapt its behaviour to match the human's 'type' [88].

These different learning approaches (sometimes unintentionally) make claims about alternative mechanisms underlying face-to-face interaction. For example, consider the concept of Theory of Mind [22], illustrated in figure 4*a*. The ability to attribute and predict mental states of others, and act in accordance with these predictions, has been argued to be one of the hallmarks of human social cognition. Figure 4*a* illustrates a 'cognitive hierarchy' [89] showing different levels of cognitive sophistication. At the lowest level, agents simply respond to features of the environment. For example, in the Prisoner's Dilemma, an agent that follows a fixed policy (e.g. splits 80% of the time and steals 20% of the time) is referred to as a Level-0 agent. Level-0 agents might differ in how they respond to the environment in what is called their 'type' in the Theory of Mind literature. For example, referring to figure 1, Level-0 agents might vary in their 'propensity to split'. A cooperative type might split 80% of the time, whereas a competitive type might split only 20% of the time. Level-1 agents have the ability to attribute a type to their partner and respond contingently based on this attribution.

These distinctions matter because research in social neuroscience shows that people exhibit sensitivity to their partner's level of Theory of Mind. For example, Yoshida and colleagues [90] had participants play a game called the Stag Hunt with synthetic agents that implemented different levels of reasoning (in the Stag Hunt, players have the opportunity to cooperate and hunt a stag for large reward or act independently to hunt rabbits). Participants adjusted their behaviour based on the estimated level of their partner's sophistication, and these estimates correlated with activity in participants' dorsolateral prefrontal cortex, suggesting that the alternative mechanisms realized in computer partners lend insight into the neural processes underlying interdependent behaviour.

Distinctions like level of Theory of Mind are routinely discussed in research on cognitive tasks like the Prisoner's

Dilemma but are almost universally absent from work on non-verbal interactive behaviours like active listening. Yet the distinctions apply to active-listening agents as well. For example, all the active-listening agents recently reviewed by Barquero and colleagues [6] produce backchannel actions based on observable features of human speakers without attributing mental state to the human speaker. This is equivalent to a Level-0 agent, and as we noted in our multi-listener work [87], listeners can have different 'types' (e.g. extroverted or introverted listeners). Experiments that study how listening agents change human behaviour make causal inferences by manipulating the listening agent's type (e.g. typically with the blunt manipulation of contingent feedback versus non-contingent feedback) and show that participants are acting as Level-1 actors (i.e. they attribute a mental state such as 'engagement' to the listener and adjust their behaviour by disclosing more and speaking more fluently).

Yet it is reasonable to posit that even silent listeners engage in higher-order Theory of Mind. For example, backchannelling has been argued to function in conversations as a signal to the speaker that the listener is attentive and thereby fosters rapport and mutual understanding. If this is indeed the function, it could benefit listeners to condition their behaviour on beliefs about the speaker: e.g. infer if the speaker believes the listener is truly listening, and attenuate or accentuate backchannels accordingly. From a theoretical perspective, several social–functional accounts of non-verbal expressions imply a role for the sender's beliefs about the receiver. Tickle-Degnen & Rosenthal's [63] work on rapport argued backchannels will be less common among close friends where common ground can often be assumed. Parkinson argued anger often intensifies until the signal is registered [91], and Leary and colleagues provide evidence that the expression of embarrassment is conditional on beliefs about observers [92]. As illustrated in figure 4*b*, Jin Joo Lee used this line of reasoning when implementing one of the few active-listening agents that realizes Theory of Mind [88]. Her models use Bayesian Theory of Mind [93] to attribute beliefs to the speaker. Specifically, the listener's model attends to the speaker's gaze to predict if the speaker believes the listener is actually listening (a Level-2 Theory of Mind), and adjusts listening behaviour accordingly. A robot listener that incorporated this model was viewed as more attentive than one without the higher-order model.

4. A partial ontology of learned-partners

Even for something as simple as modelling active-listening behaviours, there is a bewildering array of machine learning models, and if incorporated into learned-partners, participants could be presented with a range of mechanisms. For computer scientists designing these methods, these variations are often viewed as intermediate steps on the way to creating high-fidelity simulations of 'human behaviour' (as though there is a single gold standard to be achieved). For social psychologists and neuroscientists interested in using these methods, these variations present opportunities to manipulate social interaction, and the models in themselves may serve as hypothesized cognitive mechanisms. Perhaps more importantly, encouraging social scientists to understand and critique these learned models might profit their development, for example, by helping shift designers away from seeking a single gold standard and towards appreciating that different mechanisms are possible, even within the same individual, depending on the nature of the situation and beliefs about the partner.

One way to facilitate this interdisciplinary dialogue is to clarify the sometimes-unstated design choices that underlie different machine learning approaches. Several broad distinctions between models can be made. For example, do models allow continuous adjustment—e.g. reconsider behaviour every 30 ms [94]—or adjustment only at discrete event boundaries—e.g. only after a speaker completes an utterance [57] or when they produce discrete backchannel-inviting cues [88]. Another distinction concerns which inputs help determine contingent responses. For example, some models only learn contingencies to non-verbal cues [95], whereas others incorporate lexical information [96]. Some models only focus on the speaker whereas others consider how the listener's prior behaviour might have shaped the speaker's responses [57].

While all these distinctions have theoretical implications, here I emphasize two key distinctions that seem especially relevant for resolving contemporary debates on the nature of face-to-face interaction [18]. First is whether one person's non-verbal expressions are conditional on the beliefs of the partner (i.e. depend on a Theory of Mind), and second is whether non-verbal contingencies can be uncovered simply through observation or if they must be learned by an agent directly engaged in reward-seeking behaviour [12,97,98].

(a) Theory of Mind

As discussed above, learning models can differ based on whether they simply learn regularities in surface behaviours (e.g. an agent should nod or steal from their partner when certain features are observed in the environment) or whether they explicitly represent the beliefs and goals of social actors. For example, Huang's active-listening agent in figure 3 learns condition–action rules without imputing mental state to the speaker or listener. By contrast, Lee's Bayesian Theory of Mind approach in figure 4*b* hypothesizes that active listeners hold second-order beliefs about their partner (e.g. does the speaker believe I am listening to them?).

Here I shall not review the extensive literature on approaches to learning Theory of Mind models, but rather point readers to a recent and extensive review of such methods and their use in social neuroscience experiments [8]. Three points are worth raising about this research. First, alternative models can be viewed as making theoretical commitments about the nature and complexity of social mechanisms, highlighting their potential as a source of hypothetical interaction mechanisms in learned-partner experiments. Second, just because a problem can be solved with high levels of Theory of Mind does not mean that people necessarily use these mechanisms in natural interactions. For example, as discussed in §2, finding win–win solutions in negotiation is often assumed to require Level-1 reasoning: i.e. people form beliefs about their partner's negotiation priorities by exchanging information [45]. Yet other researchers have argued that win–win solutions can emerge through interaction without explicit Theory of Mind. This can occur if Level-0 actors iteratively react to the offer on the table [99,100]. Indeed, artificial intelligence (AI) negotiation agents without Theory of Mind often exceed the performance of those with this capacity [101]. One solution to this debate is to contrast these perspectives via learned-partner experiments. Third, none of the models or domains reviewed by Rusch's review of Theory of Mind models and domains considers the role of non-verbal communication, suggesting an opportunity for research, as, for example, Lee's second-order active-listening agent shows that higher-order models clearly have relevance to explaining the function of non-verbal communication in face-to-face interaction

.

(b) Learning objective: fit to existing data or maximize social reward?

Machine learning algorithms are optimization algorithms. They attempt to find a model that optimizes some objective function, such as minimizing prediction error or maximizing some measure of goal attainment, such as cumulative reward. When a learned model is used to drive the behaviour of an embodied agent, the agent will act as though it is trying to optimize this objective. In essence, the learning objective *is* the goal of the learned-partner, even if this goal is implicit. This choice touches on several central debates in theories of face-to-face interaction, such as if one can learn to interact simply by being a spectator or if acting in the world (or at least a simulation) is required [12,97].

All the active-listening agents described above adopt a spectatorial view of learning: their objective is to maximize fit to an existing corpus of interaction data (i.e. minimize error in the predicted listener responses). For example, the listening agent in figure 3 uses conditional random fields to predict backchannel probability as a function of speaker features. Though the model could be used in a social task, such as eliciting self-disclosure in a clinical interview [102], any success or failure on this task does not trickle back to shape the learned model. Similarly, Lee *et al.*'s Bayesian Theory of Mind approach, even though it argued conceptually that backchannels achieve a social goal in conversations, trained the model to fit an existing storytelling corpus [88]. Both algorithms seek to maximize fit-to-data though they differ in Theory of Mind: Huang's approach learns essentially condition–action rules without imputing mental state, whereas Lee *et al.*'s approach tries to improve fit by hypothesizing explicit beliefs and belief-update functions that would best explain the observed data (see also [103]).

Rather than optimizing fit to pre-existing data, algorithms such as multi-agent reinforcement [104] or belief learning [105], optimize models by directly interacting with other social actors (perhaps in simulation) in search of rewarding social outcomes. It is assumed that the agent is provided some reward from the environment, such as the value of a final deal in a negotiation [106], or a participant's observed disclosure when interacting with a listening agent (though some approaches allow intrinsic rewards such as satisfying curiosity, [107]). Such agents learn via exploration to discover how action sequences contribute to reward, but this is problematic in the multi-agent context as reward depends on the interaction with other interdependent partners. This is typically handled by creating simulated users (i.e. programs that approximately act like people) or self-play, where two or more agents are trained against one another [108]. Besides potentially improving performance, sometimes to superhuman levels [109], some have argued that the solutions that emerge from self-play lend insight into face-to-face interaction, such as how language evolves as a mechanism to facilitate goal achievement in social contexts [110].

Each of these approaches raises potential issues for learned-partner experiments. Learning from observation diverges theoretically from views that emphasize the importance of learning from interaction [12,97]. Practically, the learned models may not generalize to actual interactions (as when participants deviate from interaction trajectories found in the corpus). Thus, using these models for learned-actor experiments might inadvertently misalign with participants' expectations of actual contingent behaviour. Multi-agent reinforcement learning approaches may learn unnatural policies for different reasons as they typically train against other non-human actors and their behaviour depends on how the reward function is defined. As to the former issue, co-learning allows agents to discover social conventions that allow them to coordinate, but these can diverge significantly from human norms. For example, Lewis and colleagues used co-learning to discover effective negotiation tactics, but in doing so, agents discovered odd conventions such as saying 'I want' multiple times to convey stronger interest [111,112]—though emerging approaches aim to mitigate these effects [113]. An equally important concern is that designers typically optimize models to objective rewards but human decision-making involves subjective judgements [114]. For example, typically negotiation approaches reward agents for maximizing the individual financial value of the negotiated deal, but human negotiators attend to subjective concerns like how well they did compared with their partner, was the process fair, did they want to establish a long-term relationship, etc. [115,116]. Thus, the reward function constrains an agent's 'type' though most learning approaches assign a single value function unmotivated by psychological or theoretical considerations. On the positive side, the reward function becomes a source of variability that could be used to learn alternative interaction mechanisms.

(c) Discussion

Table 1 illustrates the partial ontology of learned-partner approaches through three active-listening models. Most learned-partner approaches to non-verbal communication, such as Huang *et al.* [74], use Level-0 models. By contrast, Ding and colleagues [117] developed a Level-1 model of how clinicians backchannel during a neurocognitive assessment that conditions its response frequency on the speaker's 'type', which is inferred during early stages of the interaction, and Lee *et al.* [88] designed a Level-2 model that conditions responses on whether the listener believes the speaker believes the listener is listening. All three models were trained by maximizing fit to prior observations and I am unaware of approaches that acquire active-listening behaviours by learning in the context of reward-seeking behaviour. Rather these approaches have been restricted to higher-level cognitive tasks such as learning dialogue strategies [108,118] or playing abstract games like poker [109]. The three models differ as to whether they do continuous or discrete prediction. Whereas Huang's model does continuous prediction, Ding's responds at utterance boundaries and Lee's responds to pre-defined listening cues. Only Lee's model is 'self-aware' in the sense that it considers the listener's prior behaviour when making new predictions. In terms of evaluation, Huang's model was used to generate simulated interactions that were judged by third party observers, whereas Ding's and Lee's models were tested in interaction with an embodied agent.

Table 1. Examples of how four learned embodied agents differ in design.

	Huang *et al.* [74]	Ding *et al.* [117]	Lee *et al.* [88]	Biancardi *et al.* [119]
mind level	Level-0	Level-1	Level-2	Level-2
learning obj.	fit-to-data	fit-to-data	fit-to-data	expected reward
segmentation	continuous	end-of-utterance	eliciting cue	end-of-utterance
self-aware	no	no	yes	yes
evaluation	third person with hand-crafted control	first person with human control	first person with Level-0 control	first person with Level-0 and random control

Finally, though not an active-listening agent, table 1 includes the work of Biancardi and colleagues [119] to illustrate an embodied agent that learns non-verbal signals while directly engaged in reward-seeking behaviour. Loosely based on Burgoon *et al.*'s Interaction Adaptation Theory [66], the agent learns to manage its impression while speaking. It relies on models trained offline to predict how warm or competent (in the sense defined by the stereotype content model: [120]) a human speaker finds a robot to be. Then during a live interaction, the robot is assigned an impression management strategy (e.g. appear competent) and learns how to control its non-verbal cues to convey this impression. The approach relies on the predicted beliefs of the human listener as the reward for a reinforcement learning algorithm. Self-reported impressions of participants interacting with this strategy, versus a non-contingent strategy or a random control, suggest the approach has promise.

5. Conclusion

The benefit of learned-partners for a science of face-to-face interaction ultimately depends on whether machine learning provides insight into the mechanisms that underlie interdependent behaviour. On the one hand, I have argued that alternative machine learning approaches can be seen as a potential tool for hypothesizing alternative mechanisms—by training algorithms under different processing assumptions, the resulting learned mechanisms can be seen as specific proposals for how the observed social phenomenon emerges. For example, can win–win negotiated solutions be discovered by an algorithm that only fits to observational data or does this require learning through (perhaps simulated) social interaction? On the other hand, learned-partners can address some of the methodological limitations of dyadic and non-contingent experiments for the study of interdependent behaviour. By incorporating learned mechanisms into embodied agents, participants can act with realistic and contingent simulations of human behaviour that still support controlled experiments.

To date, this project is still in its infancy. Many of the assumptions underlying machine learning approaches are not clearly articulated, especially in terms understandable for those outside the field. Within the domain of non-verbal communication, most experimentation of learned-partners focuses on a small class of machine learning methods that avoid Theory of Mind reasoning and learn as passive observers (i.e. fitting to examples of people interacting with each other) rather than learning while engaged in actual interactions. Experiments using learned-partners to study non-verbal communication typically use blunt manipulations and measures, such as whether the agent is more natural or evokes better social outcomes than a non-contingent agent. Yet as surveyed by Rusch and colleagues [8], this approach has proven fruitful in hypothesizing and testing interaction theories in more cognitive tasks, and these successes can serve as a road map for informing a science of face-to-face interaction.

This paper has highlighted how learned-partners can distort natural interactive processes, but this is certainly not the only issue surrounding the use of machine learning methods in social science research. Machine learning is sometimes heralded as a way to address the failure of many laboratory findings to generalize to real-world settings. For example, some have argued an undue focus on six theoretically posited emotional expressions impeded progress in facial expression research and that bottom-up analysis of natural interactions is an overdue antidote [121,122]. Indeed, in some settings, deep learning algorithms with minimal assumptions can 'rediscover' theoretically posited mechanisms. For example, deep convolutional neural networks have been found to process images similar to how images are processed in the human visual cortex (e.g. [123]). By analogy, there is hope that learned-partners will recover the very same solutions enacted by people, but this hope needs to be taken with considerable caution. Deep learning methods are notoriously hard to interpret and may settle on theoretically implausible mechanisms [124]. Simpler learning approaches are interpretable but often rely on theoretically derived features and representations. Together, such arguments demonstrate the need for strong partnerships between psychological and computational research. Using software to study the mind is hardly a new idea [14], but growing interdependency between computational and social research is a central feature of the emerging science of face-to-face interaction.

Data accessibility. This article has no additional data.

Conflict of interest declaration. I declare I have no competing interests.

Funding. This work is supported by the Air Force Office of Scientific Research, under grant FA9550-18-1-0182, and the Army Research Office under Cooperative Agreement Number W911NF-20-2-0053. The views and conclusions contained in this document are those of the author and should not be interpreted as representing the official policies, either expressed or implied, of the Army Research Office or the U.S. Government. The U.S. Government is authorized to reproduce and distribute reprints for Government purposes notwithstanding any copyright notation herein..

Acknowledgements. The manuscript benefited from the anonymous feedback of the reviewers and discussions with Kallirroi Georgila, Antonia Hamilton, Brian Parkinson, Tommy Parsons, David Pynadath and Stacy Marsella.

References

1. Hedström P, Ylikoski P. 2010 Causal mechanisms in the social sciences. *Annu. Rev. Sociol.* **36**, 49–67. (doi:10.1146/annurev.soc.012809.102632)
2. Blascovich J, Loomis J, Beall A, Swinth K, Hoyt C, Bailenson JN. 2002 Immersive virtual environment technology as a methodological tool for social psychology. *Psychol. Inquiry* **13**, 103–124. (doi:10.1207/S15327965PLI1302_01)
3. Gratch J, Rickel J, André E, Cassell J, Petajan E, Badler N. 2002 Creating interactive virtual humans: some assembly required. *IEEE Intell. Syst.* **17**, 54–61. (doi:10.1109/MIS.2002.1024753)
4. Pan X, Hamilton A. 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. (doi:10.1111/bjop.12290)
5. Wykowska A, Chaminade T, Cheng G. 2016 Embodied artificial agents for understanding human social cognition. *Phil. Trans. R. Soc. B* **371**, 20150375. (doi:10.1098/rstb.2015.0375)
6. Barquero G, Núñez J, Escalera S, Xu Z, Tu W-W, Guyon I, Palmero C. 2022 Didn't see that coming: a survey on non-verbal social human behavior forecasting. *Proc. Machine Learning Res.* **173,** 139–178.
7. Uc-Cetina V, Navarro-Guerrero N, Martin-Gonzalez A, Weber C, Wermter S. 2023 Survey on reinforcement learning for language processing. *Artif. Intellig. Rev.* **56**, 1543–1575. (doi:10.1007/s10462-022-10205-5)
8. Rusch T, Steixner-Kumar S, Doshi P, Spezio M, Gläscher J. 2020 Theory of mind and decision science: towards a typology of tasks and computational models. *Neuropsychologia* **146**, 107488. (doi:10.1016/j.neuropsychologia.2020.107488)
9. Johnson DW, Johnson RT. 2009 An educational psychology success story: social interdependence theory and cooperative learning. *Educ. Res.* **38**, 365–379. (doi:10.3102/0013189×09339057)
10. Bavelas JB, Coates L, Johnson T. 2000 Listeners as co-narrators. *J. Pers. Social Psychol.* **79**, 941–952. (doi:10.1037/0022-3514.79.6.941)
11. Gratch J, Lucas G. 2021 Rapport between humans and socially interactive agents. In *The handbook on socially interactive agents: 20 years of research on embodied conversational agents, intelligent virtual agents, and social robotics: methods, behavior, cognition*, vol. 1, 1st edn (eds B Lugrin, C Pelachaud, D Traum), pp. 433–462. New York, NY: ACM. (doi:10.1145/3477322.3477335)
12. Schilbach L, Timmermans B, Reddy V, Costall A, Bente G, Schlicht T, Vogeley K. 2013 Toward a second-person neuroscience. *Behav. Brain Sci.* **36**, 393–414. (doi:10.1017/S0140525X12000660)
13. Kopp S, Bergmann K. 2017 Using cognitive models to understand multimodal processes: the case for speech and gesture production. In *The handbook of multimodal-multisensor interfaces: foundations, user modeling, and common modality combinations*, vol. 1, pp. 239–276. New York, NY: ACM Books. (doi:10.1145/3015783.3015791)
14. Simon H. 1969 *The sciences of the artificial*. Cambridge, MA: MIT Press.
15. Waterman RW, Meier KJ. 1998 Principal-agent models: an expansion? *J. Public Admin. Res. Theory* **8**, 173–202. (doi:10.1093/oxfordjournals.jpart.a024377)
16. Suchman LA. 1987 *Plans and situated actions: the problem of human–machine communication*. New York, NY: Cambridge University Press.
17. Wilson M. 2002 Six views of embodied cognition. *Psychon. Bull. Rev.* **9**, 625–636. (doi:10.3758/bf03196322)
18. De Jaegher H, Di Paolo E, Adolphs R. 2016 What does the interactive brain hypothesis mean for social neuroscience? A dialogue. *Phil. Trans. R Soc. B* **371**, 20150379. (doi:10.1098/rstb.2015.0379)
19. Herschbach M. 2012 On the role of social interaction in social cognition: a mechanistic alternative to enactivism. *Phenomenol. Cogn. Sci.* **11**, 467–486. (doi:10.1007/s11097-011-9209-z)
20. Wu J, Axelrod R. 1995 How to cope with noise in the iterated prisoner's dilemma. *J. Conflict Res.* **39**, 183–189. (doi:10.1177/0022002795039001008)
21. Kelley HH, Stahelski AJ. 1970 Social interaction basis of cooperators' and competitors' beliefs about others. *J. Pers. Social Psychol.* **16**, 66. (doi:10.1037/h0029849)
22. Premack D, Woodruff G. 1978 Does the chimpanzee have a theory of mind? *Behav. Brain Sci.* **1**, 515–526. (doi:10.1017/S0140525X00076512)
23. Hoegen R, Stratou G, Gratch J. 2017 Incorporating emotion perception into opponent modeling for social dilemmas. In *Proc.16th Int. Conf. Autonomous Agents and Multiagent Systems, Sao Paulo, Brazil, 8–12 May 2017*, pp. 801–809. Liverpool, UK: IFAAMAS.
24. Niedenthal PM, Mermillod M, Maringer M, Hess U. 2010 The simulation of smiles (SIMS) mode: embodied simulation and the meaning of facial expression. *Behav. Brain Sci.* **33**, 417–480. (doi:10.1017/S0140525X10000865)
25. Waters SF, West TV, Mendes WB. 2014 Stress contagion: physiological covariation between mothers and infants. *Psychol. Sci.* **25**, 934–942. (doi:10.1177/0956797613518352)
26. van Baaren RB, Holland RW, Kawakami K, Av K. 2004 Mimicry and prosocial behavior. *Psychol. Sci.* **15**, 71–74. (doi:10.1111/j.0963-7214.2004.01501012.x)
27. Won AS, Bailenson JN, Stathatos SC, Dai W. 2014 Automatically detected nonverbal behavior predicts creativity in collaborating dyads. *J. Nonverbal Behav.* **38**, 389–408. (doi:10.1007/s10919-014-0186-0)
28. Murphy RO, Ackermann KA, Handgraaf M. 2011 Measuring social value orientation. *Judg. Decision Making* **6**, 771–781. (doi:10.1017/S1930297500004204)
29. Crivelli C, Fridlund AJ. 2018 Facial displays are tools for social influence. *Trends Cogn. Sci.* **22**, 388–399. (doi:10.1016/j.tics.2018.02.006)
30. Chambliss D, Schutt R. 2018 Causation and experimental design. In *Making sense of the social world: methods of investigation*, pp. 120–149. New York, NY: Sage Publications.
31. Kouchaki M, Smith IH. 2014 The morning morality effect: the influence of time of day on unethical behavior. *Psychol. Sci.* **25**, 95–102. (doi:10.1177/0956797613498099)
32. Preacher K, Rucker D, Hayes A. 2007 Addressing moderation mediation hypotheses: theory, methods, and prescriptions. *Multivariate Behav. Res.* **42**, 185–227. (doi:10.1080/00273170701341316)
33. Pirlott AG, MacKinnon DP. 2016 Design approaches to experimental mediation. *J. Exp. Social Psychol.* **66**, 29–38. (doi:10.1016/j.jesp.2015.09.012)
34. Preacher KJ, Hayes AF. 2004 SPSS and SAS procedures for estimating indirect effects in simple mediation models. *Behav. Res. Methods Instrum. Comput.* **36**, 717–731. (doi:10.3758/BF03206553)
35. Hutto DD. 2004 The limits of spectatorial folk psychology. *Mind Lang.* **19**, 548–573. (doi:10.1111/j.0268-1064.2004.00272.x)
36. Fujiwara K, Hoegen R, Gratch J, Dunbar NE. 2022 Synchrony facilitates altruistic decision making for non-human avatars. *Comput. Hum. Behav.* **128**, 107079. (doi:10.1016/j.chb.2021.107079)
37. Leong V, Byrne E, Clackson K, Georgieva S, Lam S, Wass S. 2017 Speaker gaze increases information coupling between infant and adult brains. *Proc. Natl Acad. Sci. USA* **114**, 13 290–13 –295. (doi:10.1073/pnas.1702493114)
38. Lanzetta JT, Englis BG. 1989 Expectations of cooperation and competition and their effects on observers' vicarious emotional responses. *J. Pers. Social Psychol.* **45**, 543–554. (doi:10.1037/0022-3514.56.4.543)
39. Tiedens LZ, Fragale AR. 2003 Power moves: complementarity in dominant and submissive nonverbal behavior. *J. Pers. Social Psychol.* **84**, 558–568. (doi:10.1037/0022-3514.84.3.558)
40. Parkinson C, Kleinbaum AM, Wheatley T. 2018 Similar neural responses predict friendship. *Nat. Commun.* **9**, 332. (doi:10.1038/s41467-017-02722-7)
41. Kashy DA, Kenny DA, Reis H, Judd C. 2000 The analysis of data from dyads and groups. In *Handbook of research methods in social and personality psychology*, vol. 38 (eds HT Reis, CM Judd), pp. 451–477. Cambridge, UK: Cambridge University Press.
42. Thorson KR, West TV, Mendes WB. 2018 Measuring physiological influence in dyads: a guide to designing, implementing, and analyzing dyadic physiological studies. *Psychol. Methods* **23**, 595–616. (doi:10.1037/met0000166)
43. Boker SM, Rotondo JL, Xu M, King K. 2002 Windowed cross-correlation and peak picking for the analysis of variability in the association between behavioral time series. *Psychol. Methods* **7**, 338–355. (doi:10.1037/1082-989X.7.3.338)

44. Van Kleef GA, De Dreu CKW, Manstead ASR. 2004 The interpersonal effects of anger and happiness in negotiations. *J. Pers. Social Psychol.* **86**, 57–76. (doi:10.1037/0022-3514.86.1.57)

45. Thompson LL. 1991 Information exchange in negotiation. *J. Exp. Social Psychol.* **27**, 161–179. (doi:10.1016/0022-1031(91)90020-7)

46. Georgescu AL, Kuzmanovic B, Santos NS, Tepest R, Bente G, Tittgemeyer M, Vogeley K. 2014 Perceiving nonverbal behavior: neural correlates of processing movement fluency and contingency in dyadic interactions. *Hum. Brain Mapp.* **35**, 1362–1378. (doi:10.1002/hbm.22259)

47. Yoshida W, Dolan RJ, Friston KJ. 2008 Game theory of mind. *PLoS Comput. Biol.* **4**, e1000254. (doi:10.1371/journal.pcbi.1000254)

48. de Melo C, Carnevale PJ, Read SJ, Gratch J. 2014 Reading people's minds from emotion expressions in interdependent decision making. *J. Pers. Social Psychol.* **106**, 73–88. (doi:10.1037/a0034251)

49. Kang S-H, Gratch J, Wang N, Watt J. 2008 Does contingency of agents' nonverbal feedback affect users' social anxiety? In *Proc. 7th Int. Joint Conf. Autonomous Agents and Multiagent Systems*, *Estoril, Portugal*, *12–16 May 2008,* vol. 1, pp. 120–127. Liverpool, UK: IFAAMAS.

50. Wilms M, Schilbach L, Pfeiffer U, Bente G, Fink GR, Vogeley K. 2010 It's in your eyes—using gaze-contingent stimuli to create truly interactive paradigms for social cognitive and affective neuroscience. *Social Cogn. Affect. Neurosci.* **5**, 98–107. (doi:10.1093/scan/nsq024)

51. Schilbach L, Wilms M, Eickhoff SB, Romanzetti S, Tepest R, Bente G, Shah NJ, Fink GR, Vogeley K. 2010 Minds made for sharing: initiating joint attention recruits reward-related neurocircuitry. *J. Cogn. Neurosci.* **22**, 2702–2715. (doi:10.1162/jocn.2009.21401)

52. Hoegen R, Schalk JVD, Lucas G, Gratch J. 2018 The impact of agent facial mimicry on social behavior in a prisoner's dilemma. In *Proc.18th Int. Conf. Intelligent Virtual Agents, Sydney, Australia, 5–8 November 2018*, pp. 275–280. New York, NY: ACM. (doi:10.1145/3267851.3267911)

53. Rosenthal R. 1966 *Experimenter effects in behavioral research*. New York, NY: Appleton-Century-Crofts.

54. Mell J, Gratch J. 2017 Grumpy and Pinocchio: the effect of language and strategy in human-agent negotiation. In *Proc. 16th Int. Conf. Autonomous Agents and Multiagent Systems, Sao Paulo, Brazil, 8–12 May 2017*, pp. 401–409. Liverpool, UK: IFAAMAS.

55. Johnson E, Roediger S, Lucas G, Gratch J. 2019 Assessing common errors students make when negotiating. In *Proc.19th Int. Conf. Intelligent Virtual Agents, Paris, France, 2–5 July 2019*, pp. 30–37. New York, NY: ACM. (doi:10.1145/3308532.3329470)

56. Boudin A, Bertrand R, Rauzy S, Ochs M, Blache P. 2021 A multimodal model for predicting conversational feedbacks. In *Proc. 24th Int. Conf. Text, Speech, and Dialogue, Olomouc, Czech Republic, 6 September 2021* (eds K Ekštein, F Pártl, M Konopík), pp. 120–127. Cham, Switzerland: Springer International Publishing. (doi:10.1007/978-3-030-83527-9_46)

57. Ishii R, Ren X, Muszynski M, Morency L-P. 2021 Multimodal and multitask approach to listener's backchannel prediction: can prediction of turn-changing and turn-management willingness improve backchannel modeling? In *Proc. 21st Int. Conf. Intelligent Virtual Agents, Japan, 14–17 September 2021*, pp. 131–138. New York, NY: ACM. (doi:10.1145/3472306.3478360)

58. Türker BB, Erzin E, Yemez Y, Sezgin TM. 2018 Audio-visual prediction of head-nod and turn-taking events in dyadic interactions. In *Proc. Conf. Int. Speech Commun. Assoc. (Interspeech), Hyderabad, India, 2–6 September 2018*, pp. 1741–1745. ISCA. (doi:10.21437/Interspeech.2018-2215)

59. Ben-Youssef A, Clavel C, Essid S. 2021 Early detection of user engagement breakdown in spontaneous human-humanoid interaction. *IEEE Trans. Affect. Comput.* **12**, 776–787. (doi:10.1109/TAFFC.2019.2898399)

60. Kindel WF, Christensen ED, Zylberberg J. 2019 Using deep learning to probe the neural code for images in primary visual cortex. *J. Vis.* **19**, 29. (doi:10.1167/19.4.29)

61. Gravano A, Hirschberg J. 2009 Backchannel-inviting cues in task-oriented dialogue. In *Proc. 10th Annu. Conf. Int. Speech Commun. Assoc., Brighton, UK, 6–10 September 2009*, pp. 1019–1022. ISCA. (doi:10.21437/Interspeech.2009-301)

62. Yngve VH. 1970 On getting a word in edgewise. In *Proc. 6th Regional Meeting of the Chicago Linguistic Society, Chicago, IL, 16–18 April 1970*, pp. 567–578. Chicago, IL: Chicago Linguistic Society.

63. Tickle-Degnen L, Rosenthal R. 1990 The nature of rapport and its nonverbal correlates. *Psychol. Inquiry* **1**, 285–293. (doi:10.1207/s15327965pli0104_1)

64. Duncan S, Franklin A, Parrill F, Welji H, Kimbara I, Webb R. 2004 Cognitive processing effects of 'social resonance' in interaction. *Proc. Annu. Meeting Cogn. Sci. Soc.* **26**(26), 16.

65. Kopp S. 2010 Social resonance and embodied coordination in face-to-face conversation with artificial interlocutors. *Speech Commun.* **52**, 587–597. (doi:10.1016/j.specom.2010.02.007)

66. Burgoon JK, Stern LA, Dillman L. 1995 *Interpersonal adaptation: dyadic interaction patterns*. Cambridge, UK: Cambridge University Press. (doi:10.1017/CBO9780511720314)

67. Levitan R, Gravano A, Hirschberg J. 2011 Entrainment in speech preceding backchannels. In *Proc. 49th Annu. Meeting Assoc. Comput. Ling.: Human Language Technologies*, vol. 2, pp. 113–117. Portland, OR: Association for Computational Linguistics. (doi:10.7916/D89Z9DCS)

68. Bernieri FJ, Rosenthal R. 1991 Interpersonal coordination: behavior matching and interactional synchrony. In *Fundamentals of nonverbal behavior* (eds RS Feldman, B Rimé), pp. 401–432. Cambridge, UK: Cambridge University Press.

69. Lakin JL, Jefferis VA, Cheng CM, Chartrand TL. 2003 Chameleon effect as social glue: evidence for the evolutionary significance of nonconscious mimicry. *J. Nonverbal Behav.* **27**, 145–162. (doi:10.1023/A:1025389814290)

70. Julien D, Brault M, Chartrand É, Bégin J. 2000 Immediacy behaviours and synchrony in satisfied and dissatisfied couples. *Can. J. Behav. Sci.* **32**, 84. (doi:10.1037/h0087103)

71. Fredrickson BL. 2016 Love: positivity resonance as a fresh, evidence-based perspective on an age-old topic. *Handb. Emot.* **4**, 847–858.

72. Abbe A, Brandon SE. 2013 The role of rapport in investigative interviewing: a review. *J. Invest. Psychol. Offender Profiling* **10**, 237–249. (doi:10.1002/jip.1386)

73. Drolet AL, Morris MW. 2000 Rapport in conflict resolution: accounting for how face-to-face contact fosters mutual cooperation in mixed-motive conflicts. *J. Exp. Social Psychol.* **36**, 26–50. (doi:10.1006/jesp.1999.1395)

74. Huang L, Morency L-P, Gratch J. 2010 Parasocial consensus sampling: combining multiple perspectives to learn virtual human behavior. In *Proc. 9th Int. Conf. Autonomous Agents and Multiagent Systems, Toronto, Canada, 10–14 May 2010*, pp. 1265–1272. Liverpool, UK: IFMAAS.

75. Huang L, Morency L-P, Gratch J. 2011 Virtual rapport 2.0. In *Proc. 10th Int. Conf. Intelligent Virtual Agents, Reykjavik, Iceland, 15–17 September 2011* (eds HH Vilhjálmsson, S Kopp, S Marsella, KR Thórisson), pp. 68–78. Berlin: Springer. (doi:10.1007/978-3-642-23974-8_8)

76. Schilbach L, Wohlschlaeger AM, Kraemer NC, Newen A, Shah NJ, Fink GR, Vogeley K. 2006 Being with virtual others: neural correlates of social interaction. *Neuropsychologia* **44**, 718–730. (doi:10.1016/j.neuropsychologia.2005.07.017)

77. de Kok I, Heylen D. 2011 The MultiLis corpus–dealing with individual differences in nonverbal listening behavior. *Toward autonomous, adaptive, and context-aware multimodal interfaces. Theoretical and practical issues* (eds A Esposito, AM Esposito, R Martone, VC Müller, G Scarpetta), pp. 362–375. Berlin, Germany: Springer. (doi:10.1007/978-3-642-18184-9_32)

78. Bailenson J, Yee N. 2005 Digital chameleons: automatic assimilation of nonverbal gestures in immersive virtual environments. *Psychol. Sci.* **16**, 814–819. (doi:10.1111/j.1467-9280.2005.01619.x)

79. Gratch J, Okhmatovskaia A, Lamothe F, Marsella S, Morales M, van der Werf RJ, Morency L-P. 2006 Virtual rapport. In *Proc. 6th Int. Conf. Intelligent Virtual Agents, Marina del Rey, CA, 21–23 August 2006*, pp. 14–27. Berlin, Germany: Springer. (doi:10.1007/11821830_2)

80. Gratch J, Wang N, Gerten J, Fast E. 2007 Creating rapport with virtual agents. In *Intelligent virtual agents. IVA 2007* (eds C Pelachaud, JC Martin, E André, G Chollet, K Karpouzis, D Pelé), pp. 125–138. Berlin, Germany: Springer. (doi:10.1007/978-3-540-74997-4_12)

81. Wang N, Gratch J. 2010 Don't just stare at me. In *Proc. 28th Annu. CHI Conf. Human Factors in Computing Systems, Atlanta, GA, 10–15 April 2010,*

pp. 1241–1250. New York, NY: ACM. (doi:10.1145/1753326.1753513)

82. Karacora B, Dehghani M, Krämer NC, Gratch J. 2012 The influence of virtual agents' gender and rapport on enhancing math performance. In *Proc. Annu. Meeting Cogn. Sci. Soc., Sapporo, Japan, 1–4 August 2012*, pp. 563–568.

83. Krämer NC, Karacora B, Lucas G, Dehghani M, Rüther G, Gratch J. 2016 Closing the gender gap in STEM with friendly male instructors? On the effects of rapport behavior and gender of a virtual agent in an instructional interaction. *Comput. Educ.* **99**, 1–13. (doi:10.1016/j.compedu.2016.04.002)

84. Jain V, Leekha M, Shah RR, Shukla J. 2021 Exploring semi-supervised learning for predicting listener backchannels. In *Proc. 2021 CHI Conf. Human Factors in Computing Systems, Yokohama, Japan, 8–13 May*, article no. 395. New York, NY: ACM. (doi:10.1145/3411764.3445449)

85. Park HW, Gelsomini M, Lee JJ, Zhu T, Breazeal C. 2017 Backchannel opportunity prediction for social robot listeners. In *Proc. 2017 IEEE Int. Conf. Robotics and Automation (ICRA), Singapore, 29 May–3 June 2017*, pp. 2308–2314. New York, NY: IEEE. (doi:10.1109/ICRA.2017.7989266)

86. Morency L-P, de Kok I, Gratch J. 2008 Predicting listener backchannels: a probabilistic multimodal approach. In *Intelligent virtual agents. IVA 2008* (eds H Prendinger, J Lester, M Ishizuka), pp. 176–190. Berlin, Germany: Springer. (doi:10.1007/978-3-540-85483-8_18)

87. Huang L, Gratch J. 2013 Explaining the variability of human nonverbal behaviors in face-to-face interaction. In *Intelligent virtual agents. IVA 2013* (eds R Aylett, B Krenn, C Pelachaud, H Shimodaira), pp. 275–284. Berlin, Germany: Springer. (doi:10.1007/978-3-642-40415-3_24)

88. Lee JJ, Sha F, Breazeal C. 2019 A Bayesian theory of mind approach to nonverbal communication. In *Proc.14th ACM/IEEE Int. Conf. Human–Robot Interaction (HRI), Daegu, South Korea, 11–14 March*, pp. 487–496. New York, NY: ACM. (doi:10.1109/HRI.2019.8673023)

89. Camerer C, Ho T, Chong K. 2003 Models of thinking, learning, and teaching in games. *Am. Econ. Rev.* **93**, 192–195. (doi:10.1257/000282803321947038)

90. Yoshida W, Seymour B, Friston KJ, Dolan RJ. 2010 Neural mechanisms of belief inference during cooperative games. *J. Neurosci.* **30**, 10 744–10 –751. (doi:10.1523/JNEUROSCI.5895-09.2010)

91. Parkinson B. 2001 Anger on and off the road. *Br. J. Psychol.* **92**, 507–526. (doi:10.1348/000712601162310)

92. Leary MR, Landel JL, Patton KM. 1996 The motivated expression of embarrassment following a self-presentational predicament. *J. Pers. Social Psychol.* **64**, 619–636. (doi:10.1111/j.1467-6494.1996.tb00524.x)

93. Tenenbaum JB, Griffiths TL, Kemp C. 2006 Theory-based Bayesian models of inductive learning and reasoning. *Trends Cogn. Sci.* **10**, 309–318. (doi:10.1016/j.tics.2006.05.009)

94. Ruede R, Müller M, Stüker S, Waibel A. 2019 Yeah, right, uh-huh: a deep learning backchannel predictor. In *Advanced social interaction with agents: Proc. 8th Int. Workshop Spoken Dialog Systems, Syracuse, Italy, 24–26 April 2019* (eds M Eskenazi, L Devillers, J Mariani), pp. 247–258. Cham, Switzerland: Springer International Publishing. (doi:10.1007/978-3-319-92108-2_25)

95. Terrell A, Mutlu B. 2012 A regression-based approach to modeling addressee backchannels. In *Proc.13th Annu. Meeting of the Special Interest Group on Discourse and Dialogue (SIGDIAL), Seoul, South Korea, 5–6 July 2012*, pp. 280–289. Stroudsburg, PA: ACL.

96. Bertrand RF, Blache PG, Espesser R, Rauzy S. 2007 Backchannels revisited from a multimodal perspective. In *Auditory–Visual Speech Processing, Hilvarenbeek, The Netherlands, 31 August–3 September*, pp. 1–5. ISCA.

97. Friston K, FitzGerald T, Rigoli F, Schwartenbeck P, O'Doherty J, Pezzulo G. 2016 Active inference and learning. *Neurosci. Biobehav. Rev.* **68**, 862–879. (doi:10.1016/j.neubiorev.2016.06.022)

98. Toma CL. 2014 Towards conceptual convergence: an examination of interpersonal adaptation. *Commun. Q.* **62**, 155–178. (doi:10.1080/01463373.2014.890116)

99. Kelley HH, Schenitzki DP. 1972 Bargaining. In *Experimental social psychology* (ed. CG McClintock), pp. 298–337. New York, NY: Holt, Rinehart, and Winston.

100. Pruitt DG, Lewis SA. 1975 Development of integrative solutions in bilateral negotiation. *J. Pers. Social Psychol.* **31**, 621–633. (doi:10.1037/0022-3514.31.4.621)

101. Baarslag T, Hendrikx M, Hindriks K, Jonker C. 2013 Predicting the performance of opponent models in automated negotiation. In *Proc. 2013 IEEE/WIC/ACM Int. Joint Conf. Web Intelligence (WI) and Intelligent Agent Technologies (IAT), Atlanta, GA, 17–20 November*, vol. 2, pp. 59–66. New York, NY: IEEE. (doi:10.1109/WI-IAT.2013.91)

102. Lucas G, Gratch J, King A, Morency L-P. 2014 It's only a computer: virtual humans increase willingness to disclose. *Comput. Hum. Behav.* **37**, 94–100. (doi:10.1016/j.chb.2014.04.043)

103. Si M, Marsella S, Pynadath D. 2005 Thespian: using multi-agent fitting to craft interactive drama. *AAMAS '05: Proc 4th Int. Joint Conf. Autonomous Agents and Multiagent Systems*, *Utrecht, The Netherlands, 25–29 July*, pp. 21–28. Liverpool, UK: IFAAMAS. (doi:10.1145/1082473.1082477)

104. Buşoniu L, Babuška R, De Schutter B. 2010 Multi-agent reinforcement learning: an overview. In *Innovations in multi-agent systems and applications - 1* (eds D Srinivasan, LC Jain), pp. 183–221. Berlin, Germany: Springer. (doi:10.1007/978-3-642-14435-6_7)

105. Feltovich N. 2000 Reinforcement-based vs. belief-based learning models in experimental asymmetric-information games. *Econometrica* **68**, 605–641. (doi:10.1111/1468-0262.00125)

106. Georgila K, Nelson C, Traum D. 2014 Single-agent vs. multi-agent techniques for concurrent reinforcement learning of negotiation dialogue policies. In *Proc. 52nd Annu. Meeting Assoc. Comput. Ling., Baltimore, MD, 22–27 June 2014*, vol. 1, pp. 500–510. Stroudsberg, PA: ACL. (doi:10.3115/v1/P14-1047)

107. Frank M, Leitner J, Stollenga M, Förster A, Schmidhuber J. 2014 Curiosity driven reinforcement learning for motion planning on humanoids. *Front. Neurorobiol.* **7**, 25. (doi:10.3389/fnbot.2013.00025)

108. Xiao G, Georgila K. 2018 A comparison of reinforcement learning methodologies in two-party and three-party negotiation dialogue. In *Proc 31st Int. Flairs Conf., Melbourne, FL, 21–23 May 2018* (eds K Brawner, V Rus), pp. 217–220. Palo Alto, CA: AAAI Press.

109. Brown N, Sandholm T. 2019 Superhuman AI for multiplayer poker. *Science* **365**, 885–890. (doi:10.1126/science.aay2400)

110. Jaques N, Lazaridou A, Hughes E, Gulcehre C, Ortega P, Strouse DJ, Leibo JZ, De Freitas N. 2019 Social influence as intrinsic motivation for multi-agent deep reinforcement learning. *Proc. 36th Int. Conf. Machine Learning, Long Beach, CA, 10–15 June 2019*, pp. 3040–3049. PMLR.

111. Kucera R. 2017 The truth behind Facebook AI inventing a new language. *Towards Data Sci.*, 7 August 2017.

112. Dauphin Y, Parikh D, Batra D. 2017 Deal or no deal? End-to-end learning of negotiation dialogues. In *Proc. 2017 Conf. Empirical Methods in Natural Language Processing, Copenhagen, Denmark, September 2017*, pp. 2443–2453. Stroudsberg, PA: ACL. (doi:10.18653/v1/D17-1259)

113. Jacob AP, Wu DJ, Farina G, Lerer A, Hu H, Bakhtin A, Andreas J, Brown N. 2022 Modeling strong and human-like gameplay with KL-regularized search. *Proc. 39th Int. Conf. Machine Learning, Baltimore, MD, 17–23 July 2022* (eds K Chaudhuri, S Jegelka, L Song, C Szepesvari, G Niu, S Sabato), pp. 9695–9728. PMLR.

114. Kahneman D, Tversky A. 1979 Prospect theory: an analysis of decision under risk. *Econometrica* **XLVII**, 263–291. (doi:10.2307/1914185)

115. Curhan JR, Elfenbein HA. 2006 What do people value when they negotiate? Mapping the domain of subjective value in negotiation. *J. Pers. Social Psychol.* **91**, 493–512. (doi:10.1037/0022-3514.91.3.493)

116. Fehr E, Schmidt KM. 1999 A theory of fairness, competition, and cooperation. *Q. J. Econ.* **114**, 817–868. (doi:10.1162/003355399556151)

117. Ding Z, Kang J, Ho TOT, Wong KH, Fung HH, Meng H, Ma X. 2022 TalkTive: a conversational agent using backchannels to engage older adults in neurocognitive disorders screening. In *Proc. ACM CHI Conf. Human Factors in Computing Systems, New Orleans, LA, 30 April–5 May 2022*, pp. 1–19. New York, NY: ACM. (doi:10.1145/3491102.3502005)

118. Shah P, Hakkani-Tur D, Liu B, Tür G. 2018 Bootstrapping a neural conversational agent with dialogue self-play, crowdsourcing and on-line reinforcement learning. In *Proc. 2018 Conf. N. Am. Ch. Assoc. Comput. Ling. Human Language*

Technologies, New Orleans, LA, 1–6 June 2018, vol. 3, pp. 41–51. Stroudsberg, PA: ACL. (doi:10.18653/v1/N18-3006)

119. Biancardi B, Dermouche S, Pelachaud C. 2021 Adaptation mechanisms in human–agent interaction: effects on user's impressions and engagement. *Front. Comput. Sci.* **3**, 696682. (doi:10.3389/fcomp.2021.696682)

120. Fiske ST, Cuddy AJ, Glick P, Xu J. 2002 A model of (often mixed) stereotype content: competence and warmth respectively follow from perceived status and competition. *J. Pers. Social Psychol.* **82**, 878. (doi:10.1037/0022-3514.82.6.878)

121. Barrett LF, Adolphs R, Marsella S, Martinez AM, Pollak SD. 2019 Emotional expressions reconsidered: challenges to inferring emotion from human facial movements. *Psychol. Sci. Public Interest* **20**, 1–68. (doi:10.1177/1529100619832930)

122. Jack RE, Crivelli C, Wheatley T. 2018 Data-driven methods to diversify knowledge of human psychology. *Trends Cogn. Sci.* **22**, 1–5. (doi:10.1016/j.tics.2017.10.002)

123. Zeman AA, Ritchie JB, Bracci S, de Beeck HO. 2020 Orthogonal representations of object shape and category in deep convolutional neural networks and human visual cortex. *Scient. Rep.* **10**, 2453. (doi:10.1038/s41598-020-59175-0)

124. Lapuschkin S, Wäldchen S, Binder A, Montavon G, Samek W, Müller K-R. 2019 Unmasking Clever Hans predictors and assessing what machines really learn. *Nat. Commun.* **10**, 1096. (doi:10.1038/s41467-019-08987-4)