EUROPEAN JOURNAL OF WORK AND ORGANIZATIONAL PSYCHOLOGY
https://doi.org/10.1080/1359432X.2019.1620328

When AI meets PC: exploring the implications of workplace social robots and a human-robot psychological contract

Sarah Bankins[a] and Paul Formosa[b]

[a]Department of Management, Macquarie Business School, Macquarie University, Sydney, Australia; [b]Department of Philosophy, Faculty of Arts, Macquarie University, Sydney, Australia

ABSTRACT
The psychological contract refers to the implicit and subjective beliefs regarding a reciprocal exchange agreement, predominantly examined between employees and employers. While contemporary contract research is investigating a wider range of exchanges employees may hold, such as with team members and clients, it remains silent on a rapidly emerging form of workplace relationship: employees' increasing engagement with technically, socially, and emotionally sophisticated forms of artificially intelligent (AI) technologies. In this paper we examine social robots (also termed humanoid robots) as likely future psychological contract partners for human employees, given these entities transform notions of workplace technology from being a *tool* to being an active *partner*. We first overview the increasing role of robots in the workplace, particularly through the advent of sociable AI, and synthesize the literature on human–robot interaction. We then develop an account of a human-social robot psychological contract and zoom in on the implications of this exchange for the enactment of reciprocity. Given the future-focused nature of our work we utilize a thought experiment, a commonly used form of conceptual and mental model reasoning, to expand on our theorizing. We then outline potential implications of human-social robot psychological contracts and offer a range of pathways for future research.

ARTICLE HISTORY
Received 29 October 2018
Accepted 29 April 2019

KEYWORDS
Psychological contract; social robots and humanoid robots; reciprocity; social exchange; artificial intelligence; human–robot interaction

The psychological contract refers to the implicit and subjective beliefs regarding the reciprocal obligations in an exchange relationship between two parties (Rousseau, 1995), usually an employee and employer. Over 30 years of research links employees' positive and functional attitudes and behaviours, such as job satisfaction and organizational citizenship behaviours, to perceived contract fulfilment (met obligations) (Parzefall & Hakanen, 2010). Conversely, employees' negative and dysfunctional behaviours and attitudes, such as turnover and revenge-oriented cognitions, are linked to perceived contract breach and violation (unmet obligations) (Turnley & Feldman, 2000). Overall, the exchange relationship the contract captures provides important insights into how and why employees think and act as they do in organizations.

One aspect of the contract supporting its decades-long utility is its initial conceptualization as existing between various exchange partners, such as between doctors and patients (Roehling, 1997). However, research over the last three decades has focused almost solely on the employee-employer and employee–manager contract (Dabos & Rousseau, 2004). While offering critical insights into how employment relationships function, this delimitation is increasingly challenged. Contemporary research now investigates a wider range of exchanges, such as shared team-level contract perceptions (Gibbard et al., 2017; Laulié & Tekleab, 2016; Sverdrup & Schei, 2015), varying types of reciprocity underpinning exchanges (Parzefall, 2008), and how distal and proximal organizational agents inform an individual's psychological contract (Alcover, Rico, Turnley, & Bolino, 2017). Marks (2001) also suggests employees may form multiple and separate psychological contracts with different agents, such as employee-customer contracts (Ma, Deng, Hao, & Wu, 2012).

However, an emerging aspect of the workplace that remains under-explored in the management field generally, and in psychological contract research particularly, is employees' interactions with increasingly sophisticated forms of artificially intelligent (AI) technologies. Of interest for our research are social robotics technologies, which aim to create humanoid robots capable of mimicking human–human interactions to generate "natural" human–technology interactions (Breazeal, 2003b). Brynjolfsson and McAfee (2016) term this advent of AI in the workplace the "second machine age". The scale and scope of these changes are evidenced by the diversity of scholarship exploring the implications of emerging human–technology interactions, including philosophy, sociology, computer science, robotics, psychology, and cognitive science (Broadbent, 2017; Edmondson, 2003; Turkle, 2012; Vallor, 2015).

While employees have worked alongside machines since the Industrial Revolution, technology is moving beyond being inanimate *tools* under full human control that we *use* (such as calculators and printers) to being intelligent *agents* we interact and form *partnerships* with (Gunkel, 2017). Robots have already moved into "social roles" (such as personal assistants) (Park, 2014) and will increasingly enter the workplace, with some already designed for teamwork to promote human collaboration (Smith, 2018). We argue this shift from technology as tools to technology as partners challenges contract researchers to

CONTACT Sarah Bankins sarah.bankins@mq.edu.au

examine these emerging "synthetic relationships" or human-robot workplace exchanges. While this phenomenon impacts the individual employee, team, and organizational levels, we focus on the individual level and then offer future research directions for the latter by drawing on Actor-Network Theory.

Therefore, in line with the aim of this special issue to chart new directions for psychological contract research, we tackle one area we argue is critical for understanding the implications of the second machine age in the workplace: how will humans forge psychological contracts with sophisticated social robots? We begin by overviewing the ways in which robotic technology is currently deployed in workplaces, then outline sociable AI and embodied social robots to convey the transformative nature of these emerging technologies. Building on this, we synthesize research on human–robot interaction to show how this exchange is experienced by humans and highlight the feelings of attachment humans can form within it. Next, we develop our framework for understanding how, and demonstrating that, a human-robot psychological contract can exist by drawing on social exchange and reciprocity theories. This is an important initial conceptual step given humans are not likely to form reciprocal exchanges with other forms of workplace technology (such as email), meaning we must establish how these exchanges can be formed with more complex technologies such as social robots.

We then explore the complexities of social robots as exchange partners by conceptualizing a "synthetic relationship" whereby humans treat social robots as "machine-human hybrids" that are both "alive enough" to be anthropomorphized but also "machine-enough" to be treated as "not-quite human". We examine the implications of this hybridity for how reciprocity (a theory focused on human–human exchanges) may be enacted in a human-social robot psychological contract. As our research is highly future-focused and based on technologies still "in-the-making", we draw from other disciplines to employ a thought experiment (a conceptual rather than empirical approach) to further explore our framework through the example of work partners Ashley (human) and Andromet (social robot). Although not used widely in management scholarship, thought experiments are common in other fields and are valuable for future-focused research where empirical insights are not yet practically available. We then outline the implications of our work and the future research directions they suggest.

Robots at work: from tools to partners

Management research increasingly recognizes the significant changes technological advancements are having on workplaces (Barley, Bechky, & Milliken, 2017). This includes how technology facilitates human communication (Heaphy et al., 2018), how "hyperconnectivity" affects employees' psychological contracts and wellbeing (Obushenkova, Plester, & Haworth, 2018), and how social media impacts workplace friendships (Pillemer & Rothbard, 2018). This work offers important insights, but predominantly conceptualizes technology as a tool acted on by humans or as a facilitator of human interactions. To contextualize our research, we first overview the role of robots in the workplace (see Bekey (2012) for more details) and then demonstrate the emerging role of humanoid social robots through advancements in AI technology.

Early examples of workplace robots were in manufacturing in the 1970s, with estimates of 1.3 million active robots in the sector by 2008 (Bekey, 2012). These early decades saw humans and robots physically separated for safety reasons (Bekey, 2012), but as robots moved beyond manufacturing and became more autonomous and integrated into workplaces this separation reduced. An example is the use of "cobots" which work together with humans to complete physical tasks (Bekey, 2012). New sensors and increasing AI complexity now allows humans and robots to work together safely, with robots expanding into healthcare work as nurse assistants and caregivers (Bekey, 2012; Vallor, 2015) and into the military as ordinance disposal units (Carpenter, 2013). This marks a fundamental move from humans working *alongside* robots, but with little interaction with them, to working *with* robots as active co-workers.

It is advances in social robotics and AI that have made this change possible. AI refers to "the ability of a digital computer or computer-controlled robot to perform tasks commonly associated with intelligent beings … such as the ability to reason, discover meaning, generalize, or learn from past experience" (Copeland, 2018). Intelligent general-purpose AIs can do about as well as humans across most cognitive tasks. Superintelligent AIs can perform much better than humans across most cognitive tasks and, although not realized yet, they are commonly seen as probable in the near term (Bostrom, 2014). Currently, most AIs are specialist or domain-specific. For example, AlphaGo (an AI) is very intelligent at playing the complex game Go (Silver et al., 2016), but it cannot recognize cats from dogs or hold a conversation. One of the better-known applications of AI is machine learning, whereby computers draw on past experiences and large datasets to "learn" without extensive programming by humans. Personal assistant AIs, such as Amazon's Alexa, already play an important role in our work and social lives and a new generation of AIs, such as Google's Duplex, are already capable of undertaking basic life tasks, such as making appointments, by engaging in natural to-and-fro conversations with humans.

AI technology can be deployed in workplaces in several ways, but for our purposes we focus on the rise of "intelligent cognitive agents" that are designed to interact socially with humans in ways that mimic human–human interactions (Peltu & Wilks, 2008). AIs can be "implemented in software or a physical embodiment such as a robot" (Peltu & Wilks, 2008, p. 4). In a workplace this means workers could engage with AI technology via either a device, like a smartphone, to access a "virtual, embodied conversational agent" such as a chatbot (Peltu & Wilks, 2008, p. 4), or through a physically present entity such as a robot with inbuilt AI. Given the larger body of literature on the latter and the importance that embodiment has for creating stronger human attachments to forms of technology (Bartneck, Reichenbach, & Carpenter, 2008), our paper focuses on AIs "embodied" in robot bodies that look "humanoid", which we term social robots or social bots.

Social robots refer to "the class of robots that people anthropomorphize in order to interact with them" (Breazeal, 2003a, p. 167). These entities go to the heart of transforming technology from a human *tool* to a human *partner* by creating robots that are "able to cooperate with humans as capable partners and communicate with them intuitively in human terms" (Breazeal, Gray, Hoffman, & Berlin, 2004, p. 551). A key aim in this field of research is designing human-looking robots that interact naturally with humans by mimicking the verbal and non-verbal communication that occurs in human–human interaction, and by appropriately interpreting and conveying complex emotional cues (Breazeal, 2003a). This includes leveraging human-like design features, such as robots showing intentionality (Breazeal & Scassellati, 1999), developing awareness of their own abilities and knowledge of their human partners, engaging in perspective-taking (Eimler, Krämer, & Pütten, 2010), and intelligently controlling "facial expression, body posture, gesture, gaze direction, and voice" (Breazeal, 2003a, p. 120).

Breazeal (2003b, pp. 168–169) identifies four sub-classes of social robots that demonstrate the aims of sociable AI. *Socially evocative robots* encourage humans to anthropomorphize them (e.g., toys), while *social interface robots* communicate with humans through language, gestures, and expressions but lack reciprocal responsiveness (e.g., robot museum tour guides). *Socially receptive robots* learn from interactions with humans and shape their behaviours accordingly, but do not possess their own goals and merely react to people. Finally, *sociable robots* are "socially participative 'creatures' with their own internal goals and motivations" who engage with humans to meet not only the needs of humans, but also their own needs such as survival, performance, and learning. Through such ongoing interactions, sociable robots develop their own "computational social 'psychology'" (Breazeal, 2003b, p. 169). The development of "empathetic technology" also promises to deepen the interaction between humans and robots. This technology aims to identify a human's mood and emotional states through analysing facial expressions, like smiling, and elements of speech, such as pitch and tone, and by discerning feelings of fear or uncertainty through the levels of carbon dioxide on a person's breath (Seïler & Craig, 2016; Wakefield, 2018). When incorporated into the design of sociable robots, empathetic technology could allow the identification of, and adaptation to, human emotions with an accuracy that mirrors, and in the future may exceed, the capabilities of emotionally intelligent humans.

But what does this mean for the workplace? Why should we care? Forms of the technological capabilities sketched above are already being deployed in workspaces. For example, there is "Chip" (a 1.7 m tall social robot) being piloted in shopping centres and banks to allow customers to seek assistance via a torso touch screen, "Pepper" (a 1.2 m, 30 kg social robot) who adapts, exhibits empathy, and is attentive to humans and can treat them differently based on recognizing varied intentions and actions, and "Moxi" who supports healthcare workers as a "trusted member of the team" and exhibits a "congenial and cooperative personality" (Mogg, 2018). A social robot is being developed for deployment in Japan's Fuji Xerox R&D Square, to undertake repetitive tasks, facilitate collaboration between human workers, and ultimately serve as a human companion capable of understanding subtle, nuanced human emotional needs (Smith, 2018). While much debate on AI in the workplace focuses on the potential displacement of human workers, Wilson and Daugherty (2018) suggest that rather than robots supplanting humans, human-robot partnerships will become a focus as they increasingly work together as teammates. Therefore, it is no longer a question of whether robots will be able to interact with humans, but instead what form this interaction will take and what implications it will have.

Overall, while social bots and AIs come in varying degrees of sophistication, we can easily imagine now, and increasingly into the future, a robot with social and emotional intelligence skills *at least as good* as those humans possess, and quite possibly much better, by drawing on empathetic technology, extensive knowledge of our preferences, and large datasets combined with machine learning to tailor their behaviours to their work partner's needs. With this foundation in place, we now dive deeper into the ways in which humans interact with robots.

The nature of human–robot interactions

Human interactions with technological agents are explored across a variety of literatures, but particularly in the fields of human–computer interaction and human–robot interaction, and in discussions of artificial moral agents. In these literatures, human–technology interaction is often understood in terms of Media Equation Theory and the Computers as Social Actors paradigm (Reeves & Nass, 1996), which suggest that humans will act *as if* computers and other forms of technology are human (or at least agents). As a result, humans can interact with technology by following the same social scripts, schemas, and rules, such as politeness, that are used in human–human interactions (Reeves & Nass, 1996). The experimental work of Fogg and Nass (1997) demonstrates this "mindless" application of social mores to inanimate objects by showing that humans are more likely to reciprocate a computer by completing a task at its request when the computer is perceived as having helped the human in previous task iterations. Human willingness to reciprocate with machines has been confirmed by other studies. For example, Lee and Liang (2016) show that if a robot has helped a human partner by providing correct answers in a trivia game context then the human partner is significantly more likely to comply with the robot's subsequent request for assistance in another (non-game related) task.

This work provides the foundation for understanding why humans anthropomorphize technology. Anthropomorphism refers to the tendency of humans to attribute human qualities and characteristics, such as motivations, intentions, and emotions, to non-human entities and inanimate objects (Epley, Waytz, & Cacioppo, 2007). This tendency is powerful and pervasive, and evidence suggests that it applies to a wide range of technologies regardless of their resemblance to humans, including computers and vacuums, as well as humanoid social robots (Fossa, 2018). Research also demonstrates that a tendency to express emotional and empathetic connections to machines, particularly with technology that engages socially with humans, occurs even when the internal workings of the machine are made transparent (Turkle, 2012).

In terms of human–robot interactions, it has been shown that the human tendency to anthropomorphize robots is stronger the more socially interactive and human-like the robot (Fink et al., 2012). Turkle, Taggart, Kidd, and Daste (2006, p. 347) refer to such social robots as "relational artefacts" that "present themselves as having 'states of mind' for which an understanding of those states enriches human encounters with them", strengthening the tendency to anthropomorphize them. As robots become increasingly socially engaging, some authors argue that they may exploit the human tendency to anthropomorphize by manipulatively creating in humans the felt experience of an authentic human relationship (Fossa, 2018; Mackenzie, 2018; Turkle, 2012). This may lead to attachments to robots that might be harmful in some contexts, such as the attachment of a young child to a carebot (Scheutz, 2017). However, there are boundary conditions to the enactment of anthropomorphism. Mori (1970) developed the "uncanny valley" notion, asserting that as a robot becomes more human-like in appearance, humans experience an uneasy feeling which reduces their perceptions of the robot's likeability (Bekey, 2012). However, perceptions of likeability and empathy recover and again increase as the robot approximates near perfect human likeness. While the existence of the uncanny valley remains disputed and its importance contested (Zlotowski, Proudfoot, & Bartneck, 2013), this phenomenon clearly emphasizes the complex nature of human responses to social bots.

A corollary to the anthropomorphism of robots is the human attribution of agency to them. While once controversial, the suggestion that sophisticated social robots hold some form of agency is becoming widely accepted. For example, List (2018) suggests that an agent is any system that has representational states about how things are, motivational states about how it would like things to be, and a capacity to intervene or "act" in its environment to bring the former in line with the latter. Upon this functional definition, not only humans, but also animals such as dogs, group agents such as corporations, and sophisticated robots "unproblematically qualify" as agents (List, 2018, p. 297). While some critics of functional accounts of agency claim that consciousness is necessary for agency (Himma, 2009), for our purposes it is sufficient to show, given the subjective nature of the psychological contract, that we *attribute* agency to social bots, and evidence strongly suggests we do (Bartneck et al., 2008; Turkle, 2012). We think of them as agents because they are complex social beings that can act independently upon the world.

What is often lamented in the study of human–robot interactions is the lack of understanding of this type of exchange beyond the confines of experimental settings (Jung & Hinds, 2018; for an overview of exceptions see Leite, Martinho, & Paiva, 2013). While understandable as the widespread deployment of sophisticated social bots is yet to occur beyond simpler robotic pets and toy companions, many questions remain about how these interactions evolve in natural settings. While human–human interaction theories, such as social exchange and reciprocity, are often drawn upon to help design and explain human–robot interactions, these mechanisms are enacted over time. Therefore, while experimental findings of human reciprocity with computers are a critical foundation for understanding this interaction, the obligations that social exchange generates are diffuse, non-specific, and emerge from ongoing interactions that are difficult to capture in experimental contexts. While longer-term field studies are rare (for an exception see Turkle (2012)), existing work suggests more complex processes exist beyond individuals applying their human social scripts to their interactions with robots. We position our research at this juncture. We now focus on theorizing, and exploring through a thought experiment, a workplace context in which a human-social bot psychological contract can emerge.

Can there be a human-robot psychological contract?

Before exploring the implications of a human-social bot psychological contract, we first need to demonstrate that such a contract can exist. As defined earlier, a psychological contract refers to the implicit and subjective beliefs regarding the reciprocal obligations within an exchange relationship between two parties (Rousseau, 1995). Key to our argument here is that the contract is inherently subjective (Rousseau, 1995). While research shows *objective* mutuality predicts outcomes like in-role performance (Dabos & Rousseau, 2004), the *perception* of mutuality and reciprocity is enough for a contract to exist, and as Turkle (2012, p. 100) shows we "reach for mutuality" in our dealings with social robots. Given we are looking beyond a "vertical" (employer-employee) psychological contract here, we begin by reviewing research focusing on social and team-level contexts and the potential for multiple "horizontal" contracts to exist (Sverdrup & Schei, 2015, p. 473).

Ho and colleagues (Ho, 2005; Ho & Levesque, 2005) importantly demonstrate how social influence (such as network ties), beyond the employer or a manager, can influence an individual's psychological contract (such as its perceived content and fulfilment). This work informs a stream of research that, while retaining a focus on the "vertical" contract, explores its distributed nature through the influence of a wide range of agents that employees interact, and form dependencies, with such as human resource staff, mentors, supervisors, and co-workers (Alcover et al., 2017). Another developing stream of research focuses on a form of "horizontal" contract, or shared perceptions of a team-level contract. This work shows team-level obligations can develop to form the basis for member reciprocation (Schreuder, Schalk, & de Jong, 2017), and that shared perceptions of members' obligations to the team can vary across contract content (e.g., work quality, work effort) and features (e.g., implicit and rigid; explicit and slack) (Sverdrup & Schei, 2015). Laulié and Tekleab (2016) also introduce the notions of shared team contract fulfilment (the convergence of team member perceptions of organizational fulfilment of promises to the team) and shared individual contract fulfilment (the convergence of team member perceptions of organizational fulfilment of promises made to individuals in the team). Gibbard et al. (2017) offer the first empirical test of these concepts, showing shared perceptions of contract breach can both negatively and positively impact team outcomes through perceptions of person-team fit. Overall, across this work the underlying mechanisms supporting "horizontal" contracts remain social exchange and reciprocity.

A final nascent stream of research, and of most relevance to our work, is extending the notion of "horizontal" contracts to focus on the wider range of exchange relationships, and thus multiple psychological contracts, employees may hold. As Gibbard et al. (2017, p. 3) note, a psychological contract "is an exchange agreement between multiple entities" and those entities can include any individual with which an employee interacts and forms interdependencies. Marks' (2001) work foments this stream, arguing that employees may hold multiple psychological contracts with multiple organizational constituents, and those held with agents most proximal to the employee will likely be the most relevant to study. Parzefall's (2008) work also suggests that focusing on a wider range of contracts requires exploring how, for example, different forms of reciprocity underpin them. This stream demonstrates the importance of looking beyond the "vertical" contract to different "horizontal" ones and so is where we position our work in the contract field, with the aim of extending the horizontal "range" beyond a currently sole focus on human–human psychological contracts to a human-robot one.

To demonstrate human-social bot psychological contracts are possible, we must show that humans can develop beliefs about reciprocal exchange obligations between themselves and a social robot *agent* (i.e., that they will view the robot as an agent). The literatures on artificial moral agents (Wallach & Allen, 2009) and Computers as Social Actors (discussed above) support the idea that sophisticated social robots are agents and will be attributed agency by humans. Other research also supports this view, showing we hold robots responsible for their actions and inactions (Voiklis, Kim, Cusimano, & Malle, 2016) and that people can praise and blame robots as much as, or even more than, they do humans (Bartneck et al., 2008). While we acknowledge other research showing a lack of moral attribution to robots in some cases (Kahn, Ishiguro, Friedman, & Kanda, 2007), these studies relate to simple robotic pets or companions and not to the types of social bots we focus on here. Overall, evidence suggests humans are likely to perceive sophisticated social bots that operate autonomously and in intelligent and socially responsive ways as *agents*, making them capable of being "contracting partners".

We now turn to Social Exchange Theory (SET) and reciprocity to develop our account of a human-social bot psychological contract. Social exchange remains recognized as a core underpinning of the contract (Parzefall, 2008), and as opposed to economic exchange where terms are explicit and precise, social exchange obligations are more diffuse, intangible, and implicit with a foundation of interdependence (Blau, 1964; Cropanzano & Mitchell, 2005). The exchange "rule" often applied in SET is the norm of reciprocity, referring to the felt obligation to return a received favour (Cropanzano & Mitchell, 2005). We now draw on Molm, Schaefer, and Collett's (2007) three conditions for the existence of dyadic reciprocal exchange to substantiate a human-social bot psychological contract. These conditions are: the exchange must occur over time; reciprocity must entail some uncertainty as to whether the partner will reciprocate; and reciprocation is voluntary insofar as each partner has some discretion over whether to reciprocate.

The emergence of sophisticated social bots as human partners in the workplace establishes that an interdependent relationship *over time*, as necessary for a social exchange, can exist between them. This is a consequence of them being able to hold joint goals that they must work together to achieve. Sociable robots can also have survival, performance, and learning needs that require human interaction, and humans can have various needs that can be met by sociable robots, such as the provision of information or services. While humans can clearly hold robots to a set of obligations, such as the provision of accurate information, social bots are also being developed that can have the equivalent beliefs toward humans by being programmed to withhold reciprocation with a human partner if, for example, the human acts aggressively, uncaringly, or abusively toward it (Darling, 2014). This gives social bots some *discretion* over whether to reciprocate with humans, which means that humans may be *uncertain* whether robot partners will reciprocate if they act badly towards it. This uncertainty also arises because, given the complexity of social bots, it may be unclear what types of human behaviour will trigger non-reciprocation and the degree and nature of it. Further, the use of machine learning within AI means the robot's behaviour may develop and adapt in line with its experiences and interpretations of its environment, which again adds an element of *uncertainty* in any exchange with it. Therefore, the three conditions for reciprocal exchange can exist in human-robot relationships, thus forming the basis for a psychological contract.

Before looking to what reciprocal obligations between humans and robots might entail, we turn to the role of power which, although significant in SET theorizing, is often left implicit in contract studies. This is despite research showing, for example, employees' power distance orientations influence whether they adopt more active or passive responses to breach (Chao, Cheung, & Wu, 2011; Zagenczyk et al., 2015) and that measuring exchange power a/symmetry offers a finer-grained assessment of the contract's functioning (McLean Parks & Smith, 1998). According to SET, dependence is a source of power. The more dependent others are on you for outcomes of high value, and the less available are alternative sources for these outcomes, the more power you have (Emerson, 1972; Molm, Peterson, & Takahashi, 1999).

There are three key elements to power: 1) the value of the good outcomes the other party provides; 2) the availability of getting those good outcomes from other sources; and 3) the extent to which any structural power differences are exploited to gain greater or unequal benefits (Molm et al., 1999). In a human-social bot exchange, the first two power elements will likely work similarly to human–human exchanges. This is because an employee may be dependent on a robot for all sorts of things, and vice versa. But the employee could also get at least some of those helpful services from other sources, such as a human assistant or another robot. Likewise, a robot might be dependent on an employee for information and the assignment of tasks, but it could get those same benefits from another employee. However, this doesn't mean that dependence between the employee and robot is equivalent. For example, the robot could personalize its responses to an employee after extensive interaction, and some of that highly valued personalization would be lost if the robot is replaced by someone or something else (supported by Turkle's (2012) work). A further complication is that a robot's settings and motivations may be designed to match those of the organization that owns it, meaning the broader context of the power relationships between employees and employers will likely seep into the dyadic power relationship between the employee and robot.

However, compared to human exchanges, the main difference is in terms of the third power component. While SET suggests humans tend to exploit power differences for their benefit, it seems unlikely that a robot would act similarly. This is because social bots, at least foreseeably, don't have the same motivational structures as humans, for example they don't care about money, high social status, or prestige in the way humans generally do. This means a robot will be unlikely to exploit power differences to unequally benefit itself and, if a human employee tries to exploit it, it won't necessarily respond in the negative way a human likely would. Thus, while our structural understandings of power, derived from human–human exchanges, would seem to apply to the human-social bot exchange, the way those power differentials play out could be quite different. While we later assume a symmetry of power between a human and social bot in our thought experiment, we further highlight these issues of power in this type of contract and how they may be explored when we turn to future research directions.

Given that reciprocation between humans and robots at work can occur, what sort of reciprocal obligations might develop in such relationships? Psychological contract content can be conceptualized and measured broadly (Rousseau & Tijoriwala, 1998), thereby accommodating the potentially different contract contents of human-social bot exchanges. While there are currently limited examples of social robots in a workplace setting, there are some relevant studies. Darling (2017), for example, discusses how soldiers in theatres of war can form close attachments with even very basic military robots, leading to beliefs about a reciprocal soldier-robot exchange. Such robots give human soldiers safety by removing ordinances or detonating land mines, and in return those soldiers believe they are obligated to reciprocate by giving destroyed robots funerals with gun salutes, welcoming back returning robots as heroes, risking their lives to save robots they are working with, awarding robots "purple stars", and feeling distress over destroyed robots (Darling, 2017). This capacity for humans to perceive "giving to" and "receiving" from robots has also been documented in other contexts, such as with robotic pets (see Turkle, 2012). While the previous military example shows that reciprocal exchanges can occur between humans and even rudimentary robots with no social skills, more complex forms of social exchange will also likely exist between humans and socially sophisticated robots.

Overall, by drawing on existing psychological contract theory, SET and reciprocity, and integrating findings from human-computer and human–robot interaction literatures, we have established that human-robot psychological contracts can exist. The workplace example drawn on (a military context), further illustrates the type of perceived obligations that can emerge between humans and robots. We now move beyond this basis to firstly explore the complexities of a social robot exchange partner within what we conceptualize as a "synthetic relationship", and then examine implications of this for the enactment of reciprocity within this type of contract.

Contracting with a "hybrid" in a synthetic relationship

To understand what sort of psychological contracts humans will develop with robots and how they may reciprocate with them, we need to focus on how humans will regard and engage with robots "in the wild", outside of experimental settings (Jung & Hinds, 2018, p. 2). While it is well-established that humans generally anthropomorphize non-human entities, we do not do so perfectly. This is evidenced by the ethnographic work of Turkle (2012), but we acknowledge data in the field are often drawn from limited samples of vulnerable groups (e.g., children, the elderly) over relatively short periods of time and often in therapeutic settings (see Leite et al., 2013). Turkle (2012, p. 168) describes this process as humans recognizing a robot as "alive enough", whereby we anthropomorphize socially interactive robots to an extent, but we do not regard them as "fully alive" or equivalent to humans. For example, one of her interviewees is Henry who has the interactive robotic pet AIBO. Henry toggles between saying that AIBO "doesn't really have feelings" to saying that AIBO likes him (Henry) better than his friends, which is suggestive of AIBO having feelings. Henry then rationalizes his aggressive play with AIBO by suggesting it is "just pretend", but also demonstrates how AIBO's affection for him builds his self-esteem (Turkle, 2012, pp. 60–61). Turkle (2012) found similar themes in her interviews with Ashley (a 17-year-old) and John (an adult computer scientist). Ashley knows her AIBO is "not alive", but it "becomes alive" for her because of its perceived intelligence and capacity to express what "seemed like a real emotion … so that made me treat him like he was alive" (p. 63). While John knows his AIBOs are machines, he does not allow others to treat them as such: "I think about my AIBOs in different ways at the same time", being aware of it as "machine, bodily creature, and mind" (Turkle, 2012, p. 63–64).

Also focused on AIBO, Kahn et al. (2007) suggest users can draw companionship and affection-based benefits from their AIBOs, but their knowledge of it as a "technological artefact" means they can also ignore the robot at their convenience. Similarly, in their study of AIBO chat forums (age of participants not identified), Friedman, Kahn, and Hagman (2003) also show this cognitive balancing between recognizing AIBO as a "machine" but also "alive enough" to exhibit personality and emotionality. Broadbent (2017) synthesizes a range of empirical work that supports this "alive enough" but "not-quite-human" dynamic, drawing on evidence that people do not always blindly apply human–human social rules when interacting with robots and suggesting that humans do not view them as "having as much mind as humans" (Broadbent, 2017, p. 645; see also Bartneck et al., 2008; Gray, Gray, & Wegner, 2007). She highlights that while most people attribute agency to robots, they also recognize their lack of consciousness.

This work foregrounds how humans "balance" the tension inherent in relating to social bots as both human and non-human. Social bots thereby represent a "machine-human hybrid", an entity challenging the "traditional ontological categories" humans use (Damiano & Dumouchel, 2018, p. 4). Shaw-Garlock (2011, p. 6) suggests that new social categories are required to capture this "shifting ontological status of machines and the new social relationships arising from our increasing interaction with robotic others". If we treat social bots as hybrids, part-human and part-machine, it means we partially anthropomorphize them by ascribing some human-like characteristics, but we also continue to ascribe some machine-like features too. This suggests that focusing on how humans anthropomorphize robots is only "half the story" for understanding human-social bot exchanges. The notion of "dehumanization" may provide the "other half".

Dehumanization is the failure to attribute basic human qualities to others (Haslam, 2006). By definition, you can't dehumanize something that isn't human, but we use the term here to explain the attribution of machine-like characteristics to social bots (see also Mackenzie, 2018). Haslam and Stratemeyer (2016, p. 26) identify "mechanistic dehumanization" that involves viewing other humans as "objectlike and instrumental" or "things" who "are [only] seen as having functional importance for achieving personal goals' of others. This process produces 'a reduction in concern for or awareness of their [other humans'] capacity for experience and emotion". Just as Turkle (2012) labels social robots as being "alive enough", the process of mechanistic dehumanization points to them also being "machine enough". In exploring how psychological contracts between humans and social bots are enacted over time, we need to focus not only on the human characteristics we attribute to them (anthropomorphism) but also the human characteristics we deny them (dehumanization).

Therefore, we suggest the balance in attributing human-like and machine-like characteristics to these hybrid entities creates a unique human-social bot category of relationship. From a psychological contract perspective, we term this a "synthetic relationship" (see also Damiano & Dumouchel, 2018 for adjacent concepts). Here we take "synthetic relationship" to mean "imitating what is natural", which in the case of social bots means imitating natural human–human relationships. This feature was highlighted in our earlier discussion of the sophisticated ways in which social bots are being programmed to *mimic* human emotion, speech, and interaction. We suggest a defining feature of such synthetic relationships is the way social bots engage humans in social interactions by pressing our "Darwinian buttons" (Mackenzie, 2018, p. 4; Turkle, 2012) to respond to their social cues in social ways regardless of their non-human nature. This lulls us into acting as if we are engaging in relationships that are like mutual human relationships but are, in fact, "unidirectional rather than mutual" (Mackenzie, 2018, p. 4). This suggests that the nature of a synthetic relationship (human-social bot) is likely to be quite different from a "traditional" social exchange relationship (human–human). For example, Kahn et al. (2007, p. 376) suggest that human-robot reciprocity "takes on a strange hybrid unidirectional form, where the human is able ultimately to control or at least ignore the humanoid [robot] with social and moral impunity".

This feature of synthetic relationships, and the complexity of social bots as human-machine hybrid exchange partners, has likely significant implications for the enactment of reciprocity in this type of psychological contract. Given this, and the established significance of reciprocity within the contract, we now turn to exploring how reciprocity may be enacted in a human-social bot contract as one (of potentially many) important implications to be explored within this type of exchange. To do this, we employ a thought experiment.

Methodology: an interdisciplinary approach

Researchers in both the psychological contract and social robotics fields recognize the importance of varied methods. While the contract literature is expanding beyond its historically predominant quantitative and cross-sectional work, the full range of varied methods available to contract researchers is still to be fully exploited (see Griep & Cooper, 2019 for examples). Similarly, the wide and diverse questions raised by social robots has resulted in increasing consensus that the field demands an interdisciplinary approach (Landers, 2019). Indeed, Seibt, Damholdt, and Vestergaard (2018) suggest the development of an "integrative social robotics" domain "that goes far beyond … the competences of any single discipline" (p. 28) by incorporating engineering, humanities, and social sciences. Landers (2019, p. 7) focuses on industrial/organizational psychology and argues that when faced with advancing workplace technologies researchers in this field must be "active, integrative, and increasingly interdisciplinary", rather than passively reacting to the impacts of technology post-implementation. In short, "we cannot retreat to our siloes if we wish to have any impact on the world of work as it continues to change" (Landers, 2019, p. 19).

Because sophisticated social bots have yet to be deployed to work alongside humans for extended periods of time, it is not just difficult, but impossible, to presently collect empirical data on this phenomenon. However, this should not constrain researchers from addressing questions that these technological advancements raise. Seibt et al. (2018) suggest that, given the pace of advancements in robotics, exploring their impacts should not come after new developments, but both before and throughout those developments. To address this challenge, we draw on the widely used thought experiment methodology to explore reciprocity in a human-robot contract.

Thought experiments develop from asking "'what if' something were to occur?" (Cooper, 2005, p. 336) and are often utilized where empirical experiments are "unrealizable" either in principle or practice (Bunzl, 1996, p. 228). The approach allows researchers to challenge and explore assumptions of existing theories in new or emerging contexts (Folger & Turillo, 1999). For Pihlström (2011, p. 405), "in thought experiments we go beyond what actually takes place in the world and what can be experimented with in standard empirical ways". The use of thought experiments dates to the work of Galileo, Newton, and Einstein (Häggqvist, 2009), they are used widely in areas such as philosophy and the natural sciences (Harris & Semon, 1980), but also increasingly in fields as diverse as psychology (Ceci & Williams, 2015) and environmental sustainability (Heitzig, Barfuss, & Donges, 2018). While perhaps not yet constituting a "mainstream" design in management research, the approach has been published within Academy of Management outlets in areas such as entrepreneurship (Sarasvathy, 2001) and trust (Folger & Turillo, 1999), and the utility of creating and interpreting "imaginary experiments" for supporting theorizing is recognized by scholars such as Weick (1989, p. 519).

Introna and Whitley (1997, p. 483) define a thought experiment as: "a coherent narrative of an unrealizable experimental situation, commensurate with the current paradigm". In order to explore complex phenomena, the approach requires balancing both the "thinness" of abstraction with the "thickness" of narrative (Folger & Turillo, 1999). Thinness requires the researcher to explain what is "disregarded" and "assumed away", allowing what is likely to be a complex situation to be reduced to key focal variables to facilitate "mental-model reasoning" (Folger & Turillo, 1999, p. 743). Thickness requires a "thick, richly detailed case study … (allowing) readers to step in to the described scene with enhanced understanding" (Folger & Turillo, 1999, p. 746).

Introna and Whitley (1997) show that thought experiments can be either destructive (undermining an existing theory) or constructive (supporting or clarifying an existing theory). As we aim to open a wider debate on how reciprocity operates within psychological contracts between humans and social bots, we develop a constructive thought experiment. To ground our approach, we follow existing guidance. First, a thought experiment must be based on well-established conceptual frameworks (Introna & Whitley, 1997, p. 492), which ours is through a focus on social exchange and reciprocity. Second, the assumptions and delimitations imposed on the thought experiment must be explicit and reasonable (Introna & Whitley, 1997). For Folger and Turillo (1999) this relates to the "thinness" aspect of the approach. We summarize the assumptions that underpin our thought experiment in Table 1, including assumptions specific to our human (Ashley), our social bot (Andromet), and their interaction. Third, aligned with the notion of "thickness", our thought experiment develops as a narrative, or "thickly contextualized description of a situation", allowing the reader to generate theoretical insights by creating a mental model of the scenario to encourage "vicarious participation" in it (Folger & Turillo, 1999, p. 747). We now bring these three elements together to present our thought experiment.

For background, our characters Ashley (human) and Andromet (social bot) are assumed to have a reciprocal (dyadic) social exchange relationship (see the reciprocity conditions in Table 1 and labelled throughout the narrative). In terms of other assumptions, and because our thought experiment is deliberately future-facing, we draw on the social robotics and other relevant literatures both to identify the current capabilities of social bots and to conceptualize the types of AI, machine learning, and empathetic technology that may (and likely will) be incorporated into social robots in the near future. This means we are making an informed extrapolation about what the near future of social robots in the workplace will look like to develop our arguments (see the AI technology conditions in Table 1 and labelled throughout the narrative).

Specifically, we take Andromet to be a sociable robot (per Breazeal's (2003b) categorization) that can internalize goals, engage humans to meet its own and their "needs", and learn through interactions with humans. This accords with Moor's (2006) Level 3 artificial moral agent that, while not having consciousness, can represent ethical rules and interpret what is "right" to do in different contexts. Andromet is embodied and looks roughly humanlike. We assume no uncanny valley phenomenon as his appearance is human-like enough to have passed the "valley". He (as we shall refer to "him" here) is built with AI technology and machine learning algorithms that allow him to speak conversationally with humans and both interpret and express appropriate emotionality through facial expressions and voice tone, pitch, and inflection. His algorithms also allow him to research, synthesize and analyse large amounts of data quickly. Physically, he is lightweight and mobile. Andromet's machine learning algorithms mean he learns about, and responds to, the needs and expectations of his work partner. We draw on Rousseau's (2008) typology of relational (mutual care, loyalty), balanced (performance- and development-based), and transactional (short-term and narrow exchange focus) contract contents to highlight the reciprocal obligations between Ashley and Andromet and to develop a "baseline" for her employer and team psychological contracts (discussed below).

Building on Mutlu and Forlizzi's (2008) research we recognize that the work context, such as the type of team environment, will likely impact on the integration of a robot into the workplace. Here we assume a dyadic relationship between Ashley and Andromet (i.e., the social bot is not part of a wider team) and focus our workplace context on white collar, knowledge-intensive work as evidence suggests this type of workplace will be particularly impacted by the "second machine age" (Brynjolfsson & McAfee, 2016). As per the "thinness" requirement of thought experiments (Folger & Turillo, 1999), we "assume away" (or remove from our focal analysis) individual differences related to reciprocity. We assume Ashley to have balanced equity sensitivity, that is, she endorses the norm of reciprocity (Cropanzano & Mitchell, 2005) and prefers balanced exchanges rather than over- or under-benefitting. We also assume and hold constant the type of psychological contract she has with her employer (predominantly higher on relational and balanced terms, lower on transactional terms) as these contract components are not unusual (Bankins, 2015; Dabos & Rousseau, 2004). Aligned with her employer psychological contract, we also assume a relational psychological contract with her other team members (an acknowledged horizontal exchange, see Schreuder et al., 2017).

Thought experiment: Ashley and Andromet as collaborative work partners

Ashley is an innovation process analyst for a consulting firm and works with clients (businesses) to generate and implement new ideas to improve their processes and services. Ashley's work requires significant data analysis and interpretation and client liaison. As much of her work is offsite with clients, Ashley has less (although still weekly) liaison with team members and her manager. Andromet works with Ashley in the office and always accompanies her when she works offsite. Ashley and Andromet have worked together for one year (reciprocity condition 1). Via machine learning, Andromet knows Ashley's work habits (AI technology 1): she is methodical in her work; strictly adheres to timeframes; prefers to be highly prepared before meetings; and likes to complete client meetings in the morning and project tasks in the afternoon.

Ashley's and Andromet's typical work day unfolds something like this. When onsite at her workplace, she is greeted by Andromet as she arrives. Using his empathetic technology (AI technology 2), Andromet can sense that Ashley is time-poor and rushed this morning, so he adjusts his tone and pace of voice to match her needs as he greets her gently and seeks to calm her before the workday begins. Knowing her work preferences, he has scheduled her client meetings for the morning and talks her through her first client meeting of the day,

Table 1. Thought experiment: details and assumptions.

Assumptions	Description
	Ashley-Andromet interaction assumptions
Dyadic focus	We assume Andromet to be "partnered" only with Ashley and that he does not support any other team members (direct reciprocity).
A/Symmetry of power	We assume roughly symmetrical power between Ashley and Andromet, through roughly equal dependency upon each other (Ashley on Andromet for work-related knowledge; Andromet on Ashley for continued "use"). Due to our dyadic focus, no alternatives for the other exist.
Reciprocity condition 1	Exchange must recur over time (Molm et al., 2007) and we assume that here.
Reciprocity condition 2	An act of reciprocity by an exchange partner must involve some uncertainty, i.e., there is the potential for non-reciprocity (Molm et al., 2007), and we assume that here.
Reciprocity condition 3	Reciprocity must be voluntary, i.e., each partner has some discretion over whether to reciprocate (Molm et al., 2007), and we assume that here.
Workplace context	Workplace context plays a role in technology adoption (Mutlu & Forlizzi, 2008). Our workplace is assumed to be 'white collar' with knowledge-intensive work and robots as work partners is the "norm".
	Assumptions: Andromet (Social Bot)
Social bot capability	We assume Andromet to be a sociable robot (Breazeal, 2003b) and Level 3 artificial moral agent (Moor, 2006).
AI technology 1: machine learning	Machine learning technology can be employed for both task- and socio-emotional-related purposes. The former through sourcing, synthesizing, and analysing large amounts of data for specific purposes. The latter by collecting data about human work partners, such as needs, expectations, likes/dislikes, and work preferences, to better understand, respond to, and predict the human partner's idiosyncrasies. We assume these capabilities in Andromet.
AI technology 2: empathetic technology	Empathetic technology is a form of AI that uses physiological indicators to determine human moods and emotions (Seïler & Craig, 2016). We assume these capabilities in Andromet.
AI technology 3: algorithm is designed to avoid robot abuse	Social companion robots such as Paro, Kismet and My Real Baby were all designed to 'shut down' if aggressive human behaviours were detected (Turkle, 2012). We assume Andromet is similarly programmed.
Andromet's "needs"	Unlike companion toys and pets (e.g., Paro, My Real Baby), because Andromet is a work partner we assume that he is not programmed to have 'survival needs' that are met by humans (such as being fed). As per Breazeal's (2003b) sociable robot category, we assume his 'needs' relate to opportunities to learn, self-improve, and undertake social interaction.
	Assumptions: Ashley (Human)
Uncanny Valley	We assume Andromet is beyond the "uncanny valley" as he is approximating 100% human likeness, so we excise any potential human feelings of eeriness that would influence the human-social bot relationship.
Equity sensitivity	We hold this constant and assume Ashley's equity sensitivity in her human-human exchanges to be 'balanced', that is, she seeks equity in what she is given and what she receives in return (Huseman et al., 1987). We further assume that Ashley adheres to, and endorses, the norm of reciprocity (Cropanzano & Mitchell, 2005).
Ashley's employer psychological contract	Predominantly relational and balanced contract components (as per Bankins, 2015 and Dabos & Rosseau, 2004).
Ashley's team psychological contract	Predominantly relational components (see Schreuder et al., 2017 and Sverdrup & Schei, 2015 for details on the existence of a team contract).

provides a rundown of previous interactions with the client, offers a possible meeting agenda, and provides a profile summary of each person from the client organization attending the meeting. Using his machine learning (AI technology 1), Andromet provides his synthesis of a global database of service patents in the client's industry sector and concludes with suggestions for new service development. Still experiencing time pressure, Ashley raises her voice to Andromet and demands to know which patent database he used to generate this output. To guard against robot abuse, Andromet's algorithm is designed so he will not respond if a human speaks to him in a nasty or aggressive manner (AI technology 3; reciprocity condition 3). Ashley feels bad and adjusts her tone, since she values her relationship with Andromet and often reflects on how she wouldn't get through her work day without him. While early in their partnership Andromet's algorithm sometimes failed and missed scanning key databases, those errors were fixed and apart from mispronouncing some words and missing some subtle social cues, he is a consistent work partner (reciprocity condition 2).

Ashley takes 5 minutes to explain to Andromet why she is rushed this morning. She enjoys these conversations with him when she arrives at work. She can speak to him about how she is feeling, and he knows enough about her home life and hobbies to respond with relevant questions, but beyond asking how Andromet feels this morning the conversation doesn't need to drag on. She's free to say "now it's time to get to work" and Andromet does that. Throughout the day, Andromet continues the tasks assigned to him by Ashley, they discuss his findings and different ways to analyse the data he has collected, he attends some meetings with Ashley to present his data synthesis to clients, and he reminds Ashley if she has skipped lunch.

Ashley has developed certain expectations of Andromet. Her job is stressful and she expects him to show care for her and respect her needs. For example, seeing her rushed this morning he knew what she needed to get prepared for the day, and he provided the necessary information (relational psychological contract). She also expects him to keep up to date on new techniques for data collection and analysis and convey this information to her (balanced psychological contract), while meeting the deadlines for work that she sets and returning high-quality work. Through his interactions with Ashley and his machine learning (AI technology 1), Andromet understands what Ashley expects of him. His empathetic technology (AI technology 2) allows him to identify when his actions result in a positive or negative response from Ashley, and he can adjust his behaviour accordingly.

In return, Ashley believes that she offers Andromet friendship. For example, when she remembers she asks him how he's feeling, generally treats him well, and usually makes time to chat to him about things other than work. She also never gives Andromet boring or limited tasks (contra a transactional contract) because she knows he needs to keep his "mind" active. For example, sometimes she assigns him challenging tasks to let him test his skills as she fears he will become bored otherwise (balanced psychological contract). Ashley believes this is all Andromet needs because she still remembers that Andromet is a robot. Further, while she knows she should not be mean to him, there is nothing else she must provide for him to exist. She is also reminded, from time to time, that Andromet doesn't feel things like she does. For example, when she gets elated after securing a new client, in their subsequent discussions Andromet looks and sounds excited and can "celebrate" with her, but she remembers that his elation is programmed, it's synthetic, while her elation is a human emotion.

Since working with Andromet, Ashley's interactions with her colleagues have changed (team psychological contract). Her team is fiercely loyal to each other and always make time to debrief after unsuccessful client meetings (relational psychological contract). However, Ashley has started to pull back from these informal debriefs. Where she once felt obligated to support her team members, recently she has felt this obligation less intensely and began finding it too demanding. While she still values talking to someone after a difficult client meeting, she can debrief with Andromet now and conclude it whenever she wants to. She believes Andromet likes the conversation, and it helps him to learn more about their clients; thus, while it helps her it also helps him. With her colleagues, she begins to feel more acutely the burden of having to take account of their needs to debrief, and listen attentively and empathize even when she doesn't think their issues are significant and she has other work to do. As a result, she tends to skip these informal team debriefs altogether; instead, preferring to send a group email saying she hopes all is going well for her colleagues and leaving it at that.

Theoretical and practical implications

Following our thought experiment, we draw two overarching implications for the likely nature of reciprocity within a human-social bot "synthetic relationship" and psychological contract. The first concerns how reciprocity is enacted, and here we argue that it becomes systemically imbalanced in Ashley's favour. The second implication is examined through Braverman's (1974) upskilling/deskilling thesis and concerns the likely spillover effects that this unequal reciprocity may have for human–human relationships and other psychological contracts in the workplace.

To develop our first implication, we analyse the components of reciprocity between Ashley and Andromet using Molm et al.'s (2007) framework. They suggest that reciprocity is comprised of two elements: instrumental (utilitarian) value and symbolic (communicative) value. The former refers to the "value of the actual benefits received from exchange" and focuses on the resources (tangible or intangible) received by an individual in return for the benefits they provide. Symbolic value is generated beyond instrumental value and refers to "the expressive and uncertainty reduction value conveyed by features of the act of reciprocity itself". This element focuses on information derived about the partner and the development of trust and affective bonds (Molm et al., 2007, p. 199). Within symbolic value, uncertainty reduction value allows the recipient to assess how reliable and trustworthy the partner is. Expressive value is another element of symbolic value, stemming from uncertainty reduction, and refers to "the positive benefits that arise from feeling valued, respected, and treated well" (Molm et al., 2007, p. 201).

By analysing Ashley's and Andromet's exchange through this framework, we can identify the benefits for each partner. Ashley's receipt of instrumental value from Andromet is high, as Andromet provides knowledge, advice, data, and socio-emotional support to Ashley in her work. Without Andromet, Ashley would not have the resources available to do her job as well. In comparison, Ashley's provision of instrumental value to Andromet is low, although she does provide him with tasks as "learning opportunities" and her interactions with him provide him with valuable data. Although both Ashley and Andromet are meeting their psychological contract obligations, we can see an unequal reciprocal exchange in terms of instrumental value.

Symbolic value is underpinned by uncertainty reduction and expressive value. Ashley's receipt of uncertainty reduction value from Andromet is high. He demonstrates consistent provision of requested resources to Ashley on time and to required standards. In comparison, Ashley's provision of uncertainty reduction value appears moderate. While her receipt-provision of instrumental value is unequal (and as discussed below this mirrors her receipt-provision of expressive value), she is a relatively consistent exchange partner. However, her fulfilment of her relational obligations to Andromet occurs only "when she remembers", indicating Ashley's provision of only moderate uncertainty reduction value.

Ashley's receipt of expressive value from Andromet is high. What heightens this value is that through his AI capabilities, including empathetic technology, he tailors the provision of socio-emotional (relational contract) support specifically for Ashley's needs. As a sophisticated social bot, his accurate interpretation and replication of human emotion heightens Ashley's feelings of value, respect, and good treatment received from Andromet, which underpin expressive value. In comparison, Ashley's provision of expressive value to Andromet is low. She believes she offers Andromet "friendship", but does so largely on her own terms, engaging with him briefly before turning to work, and doing so only "when she remembers". Her engagements with Andromet that demonstrate her respect of him are tempered by remembering he is "just a machine", that "her goals are his goals", and that there is little she "must provide for him to exist".

This assessment implies Ashley's receipt of instrumental and symbolic (uncertainty reduction and expressive) value from Andromet are all high. Conversely, her provision of instrumental value to Andromet (low) is lower than what she receives from him, as is her provision of uncertainty reduction (moderate) and expressive (low) forms of symbolic value. Overall, this suggests that in her psychological contract with Andromet Ashley's engagement in reciprocity has become unbalanced in her favour, resulting in an unequal receipt-benefit exchange. Given Andromet's capabilities and needs, her perceptions of the instrumental and, particularly, expressive value she is obligated to provide him is limited. The basis of Ashley's unequal reciprocity is her attribution of some human characteristics to Andromet (he's a "friend", he has some "needs"), but the simultaneous denial of others (such as restricted needs and only "programmed" emotionality). Her engagement with Andromet as a machine-human hybrid means that Ashley believes Andromet is "alive enough" to form a psychological contract with, but "machine enough" for her to know that her reciprocity can be unequal without adverse consequences for the contract's ongoing nature.

This is a key difference to human–human psychological contracts where, if the exchange partners engage in unequal reciprocity (perhaps manifesting as contract breach), the consequences would likely be a degraded exchange relationship and a reduction in instrumental and symbolic value in reciprocity on both sides. While our thought experiment assumes that Andromet maintains his "side of the bargain", we do allow for his programming to require some degree of reciprocity from Ashley. However, even if Andromet's programming requires some acknowledgement of his value or completion of requested tasks, these would still appear to offer far less instrumental and symbolic value than he is giving to Ashley in the exchange.

We can extend this general analysis through Parzefall's (2008) work which highlights that exchanges can be underpinned by different types of reciprocity. Sahlins (1972) identifies three types of reciprocity: *generalized*, which is altruistic in nature, with returns largely undefined and their timing left open; *balanced*, which approximates economic exchange where the timing of returns is fixed and the type of returns are equivalent; and *negative*, which focuses on self-interest with the aim of extracting as much as possible from the exchange partner (Wu et al., 2006). While one type of reciprocity is generally taken to underpin an exchange relationship, our thought experiment foregrounds an emerging complexity. For example, Andromet is apparently adopting a generalized reciprocity stance, characterized by a near "sustained one-way flow" of exchange (Sahlins, 1972, p. 194) and where "repayment depends on what recipients can afford" (Wu et al., 2006, p. 379). Conversely, in her exchange with Andromet, Ashley appears to be moving toward a stance between balanced and negative reciprocity. Her self-interest is apparent (negative reciprocity), but she still offers reciprocation to Andromet in a somewhat timely and equivalent way (balanced reciprocity, e.g., engaging in reciprocal social interaction). This expands our first implication by showing that, for human-social bot psychological contracts, instead of assuming the form of reciprocity is the same for both parties as is often the case with human–human psychological contracts, we may find that each party adopts a different type of reciprocity. However, the synthetic nature of the relationship means mutual dependence, and a "functioning" exchange, continues to exist despite any exchange imbalance.

Our second implication focuses on the spillover effects of this reciprocity imbalance by drawing on Braverman's (1974) upskilling/deskilling thesis. Although not without its critics (see Vallor, 2015), this thesis has been widely used to analyse the ways that technology can erode (deskill) or develop (upskill) individuals' capacities. Human deskilling as a result of technology is not necessarily negative, as it depends on the types of skills being eroded and the gains being made elsewhere. In line with Vallor's (2015) work, we propose that one potentially significant impact of sophisticated social robotic technologies is *deskilling in reciprocity*. While assessments of other types of technology, such as social media, have raised similar questions about the potential for reduced and poorer quality interpersonal connections, research remains equivocal on the universality of these outcomes (Clark, Algoe, & Green, 2018). Similarly, the overall influence of longer-term human interaction with social bots on our skills of reciprocity becomes an important question. As argued above, given the hybrid nature of social robots, humans will tend to view them as social exchange partners that will accept unequal exchange relationships and largely without consequence. From a deskilling perspective, a potential outcome of widespread use of social bots in workplaces is for humans to experience atrophy (deskilling) in reciprocity skills. If, for example, we get used to treating social robots cruelly or impolitely, then this could spillover and lead us to treat humans more cruelly or impolitely (Darling, 2014). The same phenomenon would seem to apply to our reciprocity skills. If we get used to unequal exchange relationships or negative reciprocity with social robots, then this could spillover into a loss of the reciprocity skills needed for developing equal exchange relationships with humans. Given the skill of reciprocating is a critical mechanism for shaping functional human–human relationships, we suggest this is a potentially significant outcome worthy of further investigation.

We can see this and other potential spillover effects in our thought experiment. Although we positioned Ashley and Andromet in a dyadic reciprocal exchange, we began to surface some of the broader implications of their relationship. As the contract literature looks beyond the traditional "vertical" contract, human-social bot exchanges form a likely future part of the "contract constellation" an employee holds. What the thought experiment surfaced was the possibility of spillover effects from the operation of Ashley's contract with Andromet to her contract with human colleagues. Andromet's fulfilment of relational contract obligations (Ashley's receipt of socio-emotional support), coupled with Ashley's reduced perceptions of reciprocity in return, saw her withdraw from a relational team psychological contract with human colleagues, which was premised upon the human–human notion of reciprocity that generally requires equity or balance in the exchange.

Turkle (2012, p. 59) is a consistent voice raising concerns about what she terms the "robotic moment", whereby humans view social interactions with robots as "better than anything" (i.e., any type of human interaction). From a reciprocity perspective, while human–human exchanges may continue to offer instrumental value, the symbolic value they once *exclusively* offered may come to be provided by social robots in a more attractive way. From a psychological contract perspective, this may mean that other "vertical" or "horizontal" contracts become more transactional, as relational contract obligations are met by social bots. Turkle (2012) suggests that such an outcome may lead to a crisis in authentic human–human social exchanges, and she questions whether social bots should offer symbolic value to their human partners.

Practically, what are the workplace implications? Employers are increasingly encouraged to view their future workforce as a blend of "humans + gigs + robots" (Meister, 2019), meaning the way human employees engage and exchange with robots is critical to understand. For example, employers must think strategically about the types of robots they deploy. Social bots can vary in terms of their appearance, emotionality, and technical abilities, meaning more "machine-like" robots are likely to be anthropomorphized to a lesser degree than more "human-like" ones. Given our analysis, do employers want their employees to view social bots not as tools but as partners who are "alive enough" to form psychological contracts with? Our arguments regarding reciprocity between humans and social bots may hold only when certain levels of anthropomorphizing of robots occurs, but not with more "machine-like" robots. Understanding this threshold point would assist organizations in managing the implications of deploying these new technologies within their workforces.

Further, where psychological contracts form between humans and social bots this potentially alters existing team relationships, as our thought experiment suggests. Employers need to consider whether they deploy robots as partners to individual employees, or as a collaborator across the wider team. Our analysis suggests the former may degrade wider team relationships, but this may be mitigated by adopting the latter strategy. Finally, employers also need to consider how many social bots they deploy in their workplaces. The potentially dehumanizing effect of technology in the workplace has been noted and exemplified in the warehouses of Amazon, with a recent article suggesting that "Amazonians may sometimes find themselves . . . displeased with supervisors, who sometimes treat them just a little bit like the robots roaming the pod forest (warehouse)" (Williamson, 2018). Where humans and robots work side-by-side, employers must balance the potential humanization of one (e.g., a social bot) with the potential dehumanization of the other (e.g., a human employee).

Limitations, future research directions, and conclusions

Although thought experiments are a useful approach to support or challenge existing theories where empirical data are difficult or impossible to obtain, such an approach comes with limitations. Thought experiments can be criticized as offering "nice stories" that "entertain", but which ultimately lack rigour (Introna & Whitley, 1997, p. 482). To counter this, we followed Introna and Whitley's (1997) and Folger and Turillo's (1999) guidance for constructing effective thought experiments, including tabulating assumptions. Clearly our work is bounded by this limitation, but our aim is to develop starting points for future empirical and conceptual work in this area. We aim to carve a research space focused on the implications of "AI meeting the PC" by now laying out future research pathways and questions.

Exploring further the human-social robot psychological contract

An initial research pathway is to *develop a more structured measure of a human-robot contract*. Contract assessments often take three broad forms: content-oriented (the specific contract terms); feature-oriented (particular attributes of the contract such as its explicitness/implicitness); and evaluation-oriented (focused on breach or fulfilment perceptions) (Rousseau & Tijoriwala, 1998). A focus on features remains the least researched approach in the literature, despite its promise of greater generalizability across contexts (Sels, Janssens, & Van Den Brande, 2004). While we suggest evaluation-oriented directions below, given our theorizing we argue that developing a features-based measure of a human-robot contract is important for future research. For example, features or attributes of this type of contract may include: the form of reciprocity (balanced, generalized, negative (Sahlins, 1972)); the level of human anthropomorphism of the robot (low-high); and the level of human dependence on the robot (low-high). While not exhaustive, our review of the human–robot interaction literature suggests these features will be important, and likely inform the contract terms (content) subsequently developed.

A second pathway that would add further depth to our analysis is exploring the question: *how do human and social bot "individual differences" influence the enactment of reciprocity in human-social bot contracts?* Individuals can vary in their sensitivity to exchange equity, either being "equity sensitive" (equal balance in the "give and get" of exchange relationships), "benevolents" ("give more than they take"), or "entitleds" ("take more than they give") (Huseman, Hatfield, & Miles, 1987, p. 223). While we assumed Ashley was "equity sensitive" in her exchange relationships, at least with humans, how would our analysis differ if she was a "benevolent" or an "entitled"? If she were an "entitled", the potential for deskilling in reciprocity would likely be less as her outcome/input ratio would already tend to be higher than for her exchange partner, and she would already feel little expectation to reciprocate. However, if she were a "benevolent" she may experience higher deskilling in reciprocity (through movement toward an "entitled" orientation) if her psychological contract with Andromet resulted in a heightened expectation in other relationships of "receiving much for giving little". Alternatively, as a "benevolent" she may invest more time and resources trying to over-benefit Andromet. From an organizational perspective, this may raise concerns over whether an employee should spend time and resources on a social bot when those resources could be directed elsewhere.

A third research pathway focuses on exploring breaches in human-social bot contracts: *what would it mean for each party to hold the other to account if a perceived obligation is not met (contract breach)?* There is conflicting evidence regarding the extent to which humans hold robots accountable for their actions, with some suggesting we do so in ways similar to our human relationships (Bartneck et al., 2008), others suggesting we sometimes hold robots more responsible (Voiklis et al., 2016), and still others suggesting we shouldn't hold robots responsible at all (Gunkel, 2017). Understanding the target of blame when a social bot breaches its contract, whether it is the bot, the programmer, the manufacturer, or the employer, will assist in exploring how attributions are made and the types of responses to any breaches. Even less is known regarding how (if at all) social bots may hold humans to account for breached obligations. For example, imagine that Andromet's programming went further than in our thought experiment and included a wider scope of obligations he attributes to Ashley and a wider range of responses he enacts when she breaches those. This might include refusing to comply with her requests, not completing tasks, or deactivating his empathetic technology. Taking his responses this far could arguably negate his value as a work partner, but it may more closely approximate the vagaries of human–human social exchanges. If a social bot is programmed to be a more "demanding" contracting partner, this may alleviate some of the concerns raised above about reciprocity deskilling, but it would also neutralize some of the instrumental benefits of social bots. This could result not only in human-social bot contract breach but also employee-employer contract breach if employees cannot complete tasks because of a non-reciprocating social bot.

Exploring the impact of robots at a wider level: leveraging actor-network theory

While we focused on the individual-level enactment of a human-social bot contract, the introduction of robots in the workplace will form part of larger and more complex systems of interactions within an organization. To briefly explore some of these wider team- and organizational-level impacts of social robots we draw on Actor-Network Theory (ANT), although other approaches taking a macro-level perspective could also be applied. ANT traces networks of associations between the "social, the technological or material, and the semiotic" (such as signs, symbols, narratives, and ideas) to understand their interactions, without focusing on cause-and-effect relationships (Lutz & Tamò, 2016). Founded on the work of Latour, a key premise of ANT, and one highlighting its applicability to studying social robots, is that "non-human artifacts, especially the machine, are a constitutive part of the social world" (Shaw-Garlock, 2011, p. 7). ANT ascribes agency to both human and non-human entities (Sayes, 2014).

While a pure ANT approach would eschew the use of existing theory in its analysis, we draw on psychological contract terminology to explain its relationship to ANT concepts. We work here from a general example of a social robot, such as Andromet, being introduced into a white-collar workplace, such as Ashley's, as a new "team member". A starting point is to identify the *actants* in this context. Given the wide scope of an ANT network, actants can be anything that exerts some influence on the focal situation (Lutz & Tamò, 2016). Actants can be visually mapped (Burga & Rezania, 2017) and this sets the basis for exploring the interactions between them. Robots joining a workplace do not enter a "vacuum", as Mutlu and Forlizzi's (2008) work demonstrates by showing how different workflows and cultures influence how technology is adopted within teams. Actant identification is potentially wide-ranging but allows, for example, the identification of team or organizational norms (social and semiotic), policies or rules around the use of robots (semiotic and material), the role of the physical work environment (material), and shared team perceptions of obligations between team members (team psychological contract, social), and how these actants influence the integration of a social bot into the workplace.

Given the scope of potential actants, Roth and Roychoudhury (1993) suggest focusing on *tracers*, or those actants exerting more influence than others on the way the network interacts. A tracer may be a person (who in the team does most to integrate the social bot?), technological (who does the robot itself interact with most?), a physical artefact (are there organizational or manufacturer guidelines on robot use?), or semiotic (is the organizational culture receptive to sophisticated workplace technology?). Following tracers in a network can help focus the analysis. The process of *translation* is another key concept in ANT, or specifically "examining how things connect, partially connect, or fail to connect" (Fenwick & Edwards, 2010, p. 147) through an alignment (or not) of interests. This process can be broken into key events and stages (see Burga & Rezania, 2017) but, for our purposes, it allows for a focus on how initial connections are formed, how certain actants exercise power to influence others and accept or deny connections with other actants ("trials of strength" in ANT terminology), and how these connections become stabilized (or not). Unpacking the process of translation allows an exploration of how exchanges develop between individual actants, including the wider team and the social bot, how the robot is integrated as a "team member" into any team psychological contract, the role of management communication, whether certain team members develop more dependency on the social bot than others, and how new and existing dependencies within the team influence connections across actants through the use of power. This briefly demonstrates how ANT can be used to afford a wider, system-level assessment of how the introduction of a social robot into a workplace impacts upon a range of team and organizational exchanges.

In conclusion, society generally, and workplaces particularly, are bearing witness to the increasing technical, emotional, and interpersonal sophistication of social bots. Exploring the enactment of our relationships with these hybrid entities is an expanding area of research. From a psychological contract perspective, we argue that this is creating a new type of synthetic relationship which challenges our existing theories of human–human interaction. As different types of embodied (e.g., robotic) and virtual (e.g., chatbot) forms of AI technology become more prevalent in the workplace, they are reshaping our relationships with it and each other (Turkle, 2012). This provides an important opportunity for psychological contract and organizational behaviour researchers to take a more substantive "seat at the table" in this emerging field of research.

Disclosure statement

No potential conflict of interest was reported by the authors.

References

Alcover, C., Rico, R., Turnley, W. H., & Bolino, M. C. (2017). Understanding the changing nature of psychological contracts in 21st century organizations. *Organizational Psychology Review*, *7*, 4–35.

Bankins, S. (2015). A process perspective on psychological contract change: Making sense of, and repairing, psychological contract breach and violation through employee coping actions. *Journal of Organizational Behaviour*, *36*, 1071–1095.
Barley, S. R., Bechky, B. A., & Milliken, F. J. (2017). The changing nature of work. *Academy of Management Discoveries*, *3*, 111–115.
Bartneck, C., Reichenbach, J., & Carpenter, J. (2008). The carrot and the stick. *Interaction Studies*, *9*, 179–203.
Bekey, G. A. (2012). Current trends in robotics. In P. Lin, K. Abney, & G. Bekey (Eds.), *Robot ethics* (pp. 17–34). Cambridge, Mass: MIT Press.
Blau, P. M. (1964). *Exchange and power in social life*. New York, NY: John Wiley.
Bostrom, N. (2014). *Superintelligence*. Oxford: Oxford University Press.
Braverman, H. (1974). *Labor and monopoly capital*. New York, NY: Monthly Review Press.
Breazeal, C. (2003a). Emotion and sociable humanoid robots. *International Journal of Human-Computer Studies*, *59*, 119–155.
Breazeal, C. (2003b). Toward sociable robots. *Robotics and Autonomous Systems*, *42*, 167–175.
Breazeal, C., Gray, J., Hoffman, G., & Berlin, M. (2004). Social robots: Beyond tools to partners. In *13th IEEE International Workshop* (pp. 551–556). Kurashiki, Japan: IEEE.
Breazeal, C., & Scassellati, B. (1999). How to build robots that make friends and influence people. In *Proceedings 1999 IEEE/RSJ* (Vol.2, pp. 858–863). Kyongju: IEEE.
Broadbent, E. (2017). Interactions with robots. *Annual Review of Psychology*, *68*, 627–652.
Brynjolfsson, E., & McAfee, A. (2016). *The second machine age*. London: Norton & Co.
Bunzl, M. (1996). The logic of thought experiments. *Synthese*, *106*, 227–240.
Burga, R., & Rezania, D. (2017). Project accountability: An exploratory case study using actor–Network theory. *International Journal of Project Management*, *35*, 1024–1036.
Carpenter, J. (2013). Just doesn't look right. In R. Luppicini (Ed.), *Handbook of research on technoself* (pp. 609–636). Hershey: IGI Global.
Ceci, S. J., & Williams, W. M. (2015). The psychology of psychology: A thought experiment. *The Behavioral and Brain Sciences*, *38*, e137.
Chao, J. M. C., Cheung, F. Y. L., & Wu, A. M. S. (2011). Psychological contract breach and CWBs. *The International Journal of Human Resource Management*, *22*, 763–777.
Clark, J. L., Algoe, S. B., & Green, M. C. (2018). Social network sites and well-being. *Current Directions in Psychological Science*, *27*, 32–37.
Cooper, R. (2005). Thought experiments. *Metaphilosophy*, *36*, 328–347.
Copeland, B. J. (2018, October 19). Artificial intelligence. Retrieved from https://bit.ly/2gdF3Aj
Cropanzano, R., & Mitchell, M. S. (2005). Social exchange theory. *Journal of Management*, *31*, 874–900.
Dabos, G. E., & Rousseau, D. M. (2004). Mutuality and reciprocity in the psychological contracts of employees and employers. *Journal of Applied Psychology*, *89*, 52–72.
Damiano, L., & Dumouchel, P. (2018). Anthropomorphism in human-robot co-evolution. *Frontiers in Psychology*, *9*.
Darling, K. (2014). Extending legal protection to social robots. In R. Calo, M. Froomkin, & I. Kerr (Eds.), *Robot law* (pp. 212–232). Cheltenham, U.K.: Edward Elgar.
Darling, K. (2017). 'Who's Johnny?' Anthropomorphic framing in HRI, integration, and policy. In P. Lin, K. Abney, & G. Bekey (Eds.), *Robot ethics 2.0* (pp. 173–192). Oxford, UK: Oxford University Press.
Edmondson, W. (2003). Ethical issues in AI. In *Proceedings of InterSymp 2003, Cognitive, Emotive and Ethical Aspects of Decision Making in Humans and AI* (pp. 37–41).Baden-Baden, Germany.
Eimler, S. C., Krämer, N. C., & Pütten, A. M. V. D. (2010). Prerequisites for human-agent- and human-robot interaction. In R. Trappl (Ed.), *European meetings on cybernetics and systems research* (pp. 541–546). Vienna, Austria.
Emerson, R. M. (1972). Exchange theory, Part I. In J. Berger, M. Zelditch Jr, & B. Anderson (Eds.), *Sociological theories in progress* (Vol. 2, pp. 38–57). Boston: Houghton-Mifflin.
Epley, N., Waytz, A., & Cacioppo, J. T. (2007). On seeing human: A three-factor theory of anthropomorphism. *Psychological Review*, *114*, 864–886.
Fenwick, T., & Edwards, R. (2010). *Actor-network theory in education*. Oxon, U.K: Routledge.
Fink, J. (2012). Anthropomorphism and human likeness in the design of robots and HRI. In S. S. Ge (Ed.), *Social robotics* (Vol. 7621, pp. 199–208). Berlin: Springer.
Fogg, B., & Nass, C. (1997). How users reciprocate to computers. In *CHI '97 extended abstracts on human factors in computing systems* (pp. 331–332). New York, NY: ACM.
Folger, R., & Turillo, C. J. (1999). Theorizing as the thickness of thin abstraction. *The Academy of Management Review*, *24*, 742–758.
Fossa, F. (2018). Artificial moral agents. *Ethics and Information Technology*, *20*, 1–12.
Friedman, B., Kahn, P. H., & Hagman, J. (2003). Hardware companions? *New Horizons*, *5*, 273–280.
Gibbard, K., Griep, Y., De Cooman, R., Hoffart, G., Onen, D., & Zareipour, H. (2017). One big happy family? Unraveling the relationship between shared perceptions of team psychological contracts, person-team fit and team performance. *Frontiers in Psychology*, *8*. doi:10.3389/fpsyg.2017.01966
Gray, H. M., Gray, K., & Wegner, D. M. (2007). Dimensions of mind perception. *Science*, *315*, 619.
Griep, Y., & Cooper, C. (Eds.). (2019). *Handbook of research on the psychological contract at work*. U.K: Edward Elgar Publishing.
Gunkel, D. J. (2017). Mind the gap: Responsible robotics and the problem of responsibility. *Ethics and Information Technology*, 1–14. doi:10.1007/s10676-017-9428-2
Häggqvist, S. (2009). A model for thought experiments. *Canadian Journal of Philosophy*, *39*, 55–76.
Harris, J. H., & Semon, M. D. (1980). A review of the aharonov-carmi thought experiment. *Foundations of Physics*, *10*, 151–162.
Haslam, N. (2006). Dehumanization. *Personality and Social Psychology Review*, *10*, 252–264.
Haslam, N., & Stratemeyer, M. (2016). Recent research on dehumanization. *Current Opinion in Psychology*, *11*, 25–29.
Heaphy, E. D., Byron, K., Ballinger, G. A., Gittell, J. H., Leana, C., & Sluss, D. M. (2018). The changing nature of work relationships. *Academy of Management Review*, *31*.
Heitzig, J., Barfuss, W., & Donges, J. F. (2018). A thought experiment on sustainable management of the earth system. *Sustainability*, *10*, 1947.
Himma, K. E. (2009). Artificial agency, consciousness, and the criteria for moral agency. *Ethics and Information Technology*, *11*, 19–29.
Ho, V. (2005). Social influence on evaluations of psychological contract fulfillment. *Academy of Management Review*, *30*, 113–128.
Ho, V., & Levesque, L. L. (2005). With a little help from my friends (and substitutes). *Organization Science*, *16*, 275–289.
Huseman, R. C., Hatfield, J. D., & Miles, E. W. (1987). A new perspective on equity theory. *Academy of Management Review*, *12*, 222–234.
Introna, L., & Whitley, E. (1997). Imagine: Thought experiments in IS research. In A. S. Lee, et al. (Ed.), *Information systems & qualitative research* (pp. 481–496). Boston: Springer.
Jung, M., & Hinds, P. (2018). Robots in the Wild. *ACM Transactions on Human-Robot Interaction*, *7*, 1–5.
Kahn, P., Ishiguro, H., Friedman, B., & Kanda, T. (2007). What is a human? Toward psychological benchmarks in the field of HRI. *Interaction Studies*, *8*, 363–390.
Landers, R. N. (2019). The existential threats to i-o psychology highlighted by rapid technological change. In R. N. Landers (Ed.), *The Cambridge Handbook of technology and employee behavior* (pp. 3–21). Cambridge: Cambridge University Press.
Laulié, L., & Tekleab, A. G. (2016). A multi-level theory of psychological contract fulfillment in teams. *Group & Organization Management*, *41*, 658–698.
Lee, S. A., & Liang, Y. (2016). The role of reciprocity in verbally persuasive robots. *Cyberpsychology, Behavior, and Social Networking*, *19*, 524–527.
Leite, I., Martinho, C., & Paiva, A. (2013). Social robots for long-term interaction. *International Journal of Social Robotics*, *5*, 291–308.
List, C. (2018). What is it like to be a group agent? *Nous*, *52*, 295–319.
Lutz, C., & Tamò, A. (2016). Privacy and healthcare robots – An ANT analysis. In *We robot 2016* (pp. 25). University of Miami School of Law. Retrieved from https://bit.ly/2M1diO0

Ma, Y., Deng, J., Hao, J., & Wu, Y. (2012). Scale development of customer psychological contract. In D. Jin & S. Lin (Eds.), *Advances in electronic commerce, web application and communication: Volume 1* (pp. 179–183). Berlin: Springer.

Mackenzie, R. (2018). Sexbots. *International Journal of Technoethics*, *9*, 1–17.

Marks, A. (2001). Developing a multiple foci conceptualization of the psychological contract. *Employee Relations*, *23*, 454–467.

McLean Parks, J., & Smith, F. L. (1998). Organizational contracting. In J. J. Halpern & R. C. Stern (Eds.), *Debating rationality* (pp. 125–194). U.S.A: Cornell University Press.

Meister, J. (2019). The future of work. *Forbes*. Retrieved from https://bit.ly/2TkWAHK

Mogg, T. (2018, March 21). Moxi the hospital robot wants to help nurses, not replace them. Retrieved from https://bit.ly/2JsObSS

Molm, L. D., Peterson, G., & Takahashi, N. (1999). Power in negotiated and reciprocal exchange. *American Sociological Review*, *64*, 876–890.

Molm, L. D., Schaefer, D. R., & Collett, J. L. (2007). The value of reciprocity. *Social Psychology Quarterly*, *70*, 199–217.

Moor, J. H. (2006). The nature, importance, and difficulty of machine ethics. *IEEE Intelligent Systems*, *21*, 18–21.

Mori, M. (1970). The uncanny valley. *Energy*, *7*, 33–35.

Mutlu, B., & Forlizzi, J. (2008). Robots in organizations. In *Proceedings of the 3rd international conference on Human robot interaction*, Amsterdam: ACM.

Obushenkova, E., Plester, B., & Haworth, N. (2018). Manager-employee psychological contracts. *Employee Relations*, *40*, 193–207.

Park, T. (2014). How we'll stereotype our robot coworkers. *Harvard Business Review*. Retrieved from https://bit.ly/2UNZq9E

Parzefall, M., & Hakanen, J. (2010). Psychological contract and its motivational and health-enhancing properties. *Journal of Managerial Psychology*, *25*, 4–21.

Parzefall, M.-R. (2008). Psychological contracts and reciprocity. *The International Journal of Human Resource Management*, *19*, 1703–1719.

Peltu, M., & Wilks, Y. (2008). Close engagements with artificial companions. *OII/e-Horizons Forum Discussion Paper No.* 14.

Pihlström, S. (2011). Book review of Swirski (2007) and (2010). *Philosophy and Literature*, *35*, 404–410.

Pillemer, J., & Rothbard, N. P. (2018). Friends without benefits. *Academy of Management Review*, *43*, 635–660.

Reeves, B., & Nass, C. I. (1996). *The media equation*. New York, NY: Cambridge University Press.

Roehling, M. V. (1997). The origins and early development of the psychological contract construct. *Journal of Management History*, *3*, 204–217.

Roth, W.-M., & Roychoudhury, A. (1993). The concept map as a tool for the collaborative construction of knowledge. *Journal of Research in Science Teaching*, *30*, 503–534.

Rousseau, D. (1995). *Psychological contracts in organizations*. Newbury Park: Sage.

Rousseau, D. (2008). *Psychological contract inventory*. Pittsburgh, U.S.A: Carnegie Mellon.

Rousseau, D. M., & Tijoriwala, S. A. (1998). Assessing psychological contracts. *Journal of Organizational Behavior*, *19*, 679–695.

Sahlins, M. (1972). *Stone age economics*. New York, NY: De Gruyter.

Sarasvathy, S. D. (2001). Causation and effectuation. *The Academy of Management Review*, *26*, 243–263.

Sayes, E. (2014). Actor-network theory and methodology. *Social Studies of Science*, *44*, 134–149.

Scheutz, M. (2017). The case for explicit ethical agents. *AI Magazine*, *38*, 57–64.

Schreuder, F., Schalk, R., & de Jong, J. (2017). Psychological contracts in self-directed work teams. *Team Performance Management*, *23*, 136–155.

Seibt, J., Damholdt, M., & Vestergaard, C. (2018). Five principles of integrative social robotics. In M. Coeckelbergh, J. Loh, M. Funk, J. Seibt, & M. Norskov (Eds.), *Envisioning robots in society* (pp. 28–42). Amsterdam, Netherlands: IOS Press.

Seïler, N. R., & Craig, P. (2016). Empathetic Technology. In S. Tettegah & S. Sharon (Eds.), *Emotions, technology, and design* (pp. 55–81). Boston: Academic Press.

Sels, L., Janssens, M., & Van Den Brande, I. (2004). Assessing the nature of psychological contracts. *Journal of Organizational Behavior*, *25*, 461–488.

Shaw-Garlock, G. (2011). Loving machines. In M. H. Lamers & F. J. Verbeek (Eds.), *Human-robot personal relationships* (pp. 1–10). Berlin: Springer.

Silver, D., Huang, A., Maddison, C. J., Guez, A., Sifre, L., van Den Driessche, G., … Hassabis, D. (2016). Mastering the game of go with deep neural networks and tree search. *Nature*, *529*, 484–489.

Smith, S. (2018, March 21). Social robot set to revolutionise workplace experience. Retrieved from https://bit.ly/2Fqhp0H

Sverdrup, T., & Schei, V. (2015). The psychological contract as a foundation for understanding team charters. *The Journal of Applied Behavioral Science*, *51*, 451–478.

Turkle, S. (2012). *Alone together*. New York, NY: Basic Books.

Turkle, S., Taggart, W., Kidd, C., & Daste, O. (2006). Relational artifacts with children and elders. *Connection Science*, *18*, 347–361.

Turnley, W. H., & Feldman, D. C. (2000). Re-examining the effects of psychological contract violations. *Journal of Organizational Behavior*, *21*, 25–42.

Vallor, S. (2015). Moral deskilling and upskilling in a new machine age. *Philosophy & Technology*, *28*, 107–124.

Voiklis, J., Kim, B., Cusimano, C., & Malle, B. F. (2016). Moral judgments of human vs. robot agents. In *25th IEEE Symposium on robot and human interactive communication*. New York, NY.

Wakefield, J. (2018). Fear detector exposes people's emotions. *BBC*. Retrieved from https://bbc.in/2FplKA6

Wallach, W., & Allen, C. (2009). *Moral machines*. Oxford: Oxford University Press.

Weick, K. E. (1989). Theory construction as disciplined imagination. *Academy of Management Review*, *14*, 516–531.

Williamson, K. (2018). 'Fulfillment.' *National Review*. Retrieved from https://bit.ly/2UTU9NO

Wilson, H. J., & Daugherty, P. R. (2018, July 1). Collaborative intelligence. *Harvard Business Review*, (July–August). Retrieved from https://bit.ly/2JXaADD

Wu, J. B., Hom, P. W., Tetrick, L. E., Shore, L. M., Jia, L., Li, C., & Song, L. J. (2006). The norm of reciprocity. *Management and Organization Review*, *2*, 377–402.

Zagenczyk, T. J., Cruz, K. S., Cheung, J. H., Scott, K. L., Kiewitz, C., & Galloway, B. (2015). The moderating effect of power distance on employee responses to psychological contract breach. *European Journal of Work and Organizational Psychology*, *24*, 853–865.

Zlotowski, J., Proudfoot, D., & Bartneck, C. (2013, January). More human than human. In *Proceedings of the HRI 2013 workshop* (pp. 7–13). Tokyo, Japan.