EMBODYING EXPERTISE
AS A PERFORMER AND
PERCEIVER

Insights from the arts and robotics

Emily S. Cross

22.1 Linking action with perception

The ability to use perceptual information about another individual’s or one’s own movements to inform subsequent movements is essential for successful interactions with the environment, and thus, for survival. Consequently, it is no surprise that inquiry into how the human brain negotiates the path between action execution and perception has intrigued venerable thinkers throughout the millennia, including Aristotle, Descartes, and William James. Until the mid-19th century, prevalent thinking on this relationship was dominated by the Cartesian sensorimotor view, which advanced the idea of independence between the perceptual and production domains (Descartes 1664, as discussed in Prinz 1997). As knowledge about the capabilities of the human brain continued to accrue, pioneering scholars began to explore the idea that perceptual processes might indeed converge or overlap to some extent with motor performance (James 1890). Empirical investigation into the brain’s ability to use perceptual information to shape movement began in the mid-20th century, during which time information processing explanations proposing complex transformations from perception to the organization and execution of action gathered momentum (Sanders 1967, 1983; Welford 1968; Massaro and Friedman 1990).

In the 1990s, research into the intersection of action perception and production experienced an extraordinary renaissance that was due in large part to the discovery of mirror neurons in the ventral premotor cortex of the monkey brain. Remarkably, these neurons discharge in a similar manner both when a monkey performs an action and when it observes another monkey or human perform the same action (di Pellegrino et al. 1992; Gallese et al. 1996; Rizzolatti et al. 1996a). As such, these specialized neurons have prompted these researchers and others to propose that action perception and production processes form a bidirectional, interactive loop within the primate brain, and that action recognition might be explained by the observer’s brain simulating the observed movements of another individual (Fadiga et al. 1995, 1999; Grafton et al. 1996; Rizzolatti et al. 2001).

This hypothesis has sparked lively debate among researchers regarding the specific parameters, scope, and limitations of such an action simulation system. From this debate, multiple theories
have emerged that are motivated by action-perception interactions, such as simulation theory (Gallese and Goldman 1998), the dual route theory (Tessari and Rumiati 2007), and the action hierarchy theory (Grafton and Hamilton 2007). However, many of these theories do not make specific predictions and consequently collapse under empirical scrutiny. The theory that has been the subject of the most empirical investigation over the past several decades is simulation theory (sometimes referred to as the direct matching hypothesis). Simulation theory maintains that human sensorimotor cortices become active under a broad array of action-related activities, ranging from executing planned, overt movements to observing or imagining actions that are never executed. According to this theory, we represent another individual’s actions by matching their movements with resonating, covert movements of our own (Gallese and Goldman 1998; Jeannerod 2001; Goldman, 2005), perhaps by using mirror neurons (Gallese and Goldman 1998; Goldman 2005). A striking and consistent result from neuroimaging studies (Grafton et al. 1996; Rizzolatti et al. 1996a; Iacoboni et al. 1999, 2005; Buccino et al. 2001; Grèzes and Decety 2001; Grèzes et al. 2001), and transcranial magnetic stimulation studies (Fadiga et al. 1995, 1999; Strafella and Paus 2000; Gangitano et al. 2001; Patuzzo et al. 2003) is that motor and premotor areas that are classically associated with movement preparation are also active when simply observing the actions of others.

Behavioural psychophysics studies have demonstrated interactions between action perception and execution (Brass et al. 2000, 2001a, 2001b; Kilner et al. 2003; Hamilton et al. 2004, 2005) and lend additional credence to the notion of overlapping neural processes for action observation and execution. However, although quantifiable overlap exists between different action activities, there is not a 1-to-1 correspondence between primary motor, premotor, and parietal cortical regions active during action observation, simulation, and execution. This is clear from meta-analyses of functional neuroimaging studies on shared representations across these action domains (Grèzes and Decety 2001).

Important for considerations of the role played by expertise in shaping the relationship between action and perception, a wealth of literature demonstrates that the more familiar an action is, the stronger the response is within these core sensorimotor regions (Buccino et al. 2004; Calvo–Merino et al. 2005; Cross et al. 2006; Shimada 2010; Gardner et al. 2015). These studies and others thus lend support to experience-driven simulation accounts of action perception (Sinigaglia 2013), which form the foundation of the direct matching hypothesis of action understanding (Gallese and Goldman 1998; Rizzolatti et al. 2001; Wolpert et al. 2003; although see Csibra 2005 and Kilner 2011 for alternative accounts). In terms of familiarity (for which expertise can be considered an extreme example), a linear relationship between magnitude of sensorimotor activity and familiarity would be consistent with this hypothesis: as familiarity or expertise increases, the simulation of how an action might unfold over time becomes more accurate and resonance between an observer’s motor system and an observed action is maximized.

While this linear relationship between an observer’s expertise or familiarity with an action and sensorimotor cortical activity has found support within the empirical literature and certainly has intuitive appeal, an increasing number of studies report findings demonstrating that this relationship is likely not that straightforward (Gazzola et al. 2007; Cross et al. 2012; Liew et al. 2013; Tipper et al. 2015; Gardner et al. 2017). These studies have all reported greater activity within sensorimotor cortices when participants observe actions that are unfamiliar or with which they have very little expertise (compared to more familiar actions), a finding that appears at odds with a simulation-based account of how the brain negotiates the link between action and perception. The findings from these studies suggest that a linear relationship between sensorimotor activity and familiarity is likely too simplistic. In terms of the direct matching
hypothesis; this theory would struggle to explain why an \textit{unfamiliar} action that is \textit{not} in the observer's repertoire would elicit greater action observation network (AON) activity. Predictive coding models of AON function (Keysers and Perrett 2004; Kilner et al. 2007a, 2007b; Gazzola and Keysers 2009; Schippers and Keysers 2011; Tipper et al. 2015), predicated on the use of perceptuomotor maps to predict and interpret observed actions (Lamm et al. 2007; Schubotz 2007; Urgesi et al. 2010), may potentially help to resolve these seemingly discrepant findings concerning the relationship between familiarity of an observed movement and engagement of sensorimotor cortices (cf. Gardner et al. 2017).

In order to explore some of the theoretical ideas introduced above in more detail, the following two sections highlight a number of empirical studies from two distinct but complementary avenues. First, I discuss work performed with expert and novice dancers, and then I examine how our understanding of a system that has evolved to code action and perception information within a common space might be co-opted for skillful interactions with artificial agents. Together, these examples should reinforce how a better understanding of the neurocognitive mechanisms and consequences of expertise help us to navigate a complex social environment in a skillful manner.

22.2 Expertise, embodiment, and the performing arts

In the past decade and a half or so, the number of researchers using performing arts experts or contexts to explore the relationship between expertise, embodiment, and perception has continued to grow. The empirical work in this domain has primarily followed one of two routes: either recruiting and examining performing arts experts (such as dancers) who have built up motoric expertise over many years of training or recruiting naive participants and training them how to perform a particular set of movements and examining the mechanisms and consequences of de novo expertise. In a chapter co-authored by Beatriz Calvo-Merino and myself, we examine in detail a broad range of studies that fit into these two categories (Cross and Calvo-Merino 2016). Rather than simply repeating what we have already covered in this previous chapter, in the following I present a brief overview of major studies and seminal findings that fit within these two primary domains, and then include some considerations regarding the relationship between expertise, embodiment, and aesthetic preference, another research area that is rapidly gaining momentum (cf. Kirsch et al. 2018).

22.2.1 Longstanding expertise

The very first study to draw upon expertise from the performing arts domain to examine how years of practice to hone a particular motor vocabulary shapes how we see the world was run by Beatriz Calvo-Merino and colleagues (Calvo-Merino et al. 2005). In this study, the authors recruited a sample of male participants from three groups: one group comprised professional ballet dancers, another group comprised expert capoeira\textsuperscript{1} performers, and a third group included men who had no particular expertise in either ballet or capoeira. All participants underwent functional magnetic resonance imaging (fMRI), a type of brain scanning that requires participants to lie still in a supine position with the head immobilized, while watching and/or listening to stimuli of the experimenter's choosing via fMRI compatible audiovisual equipment. In this study, participants watched short, silent film clips that featured an expert male ballet dancer performing a number of ballet movements and an expert capoeira performer performing capoeira movements. Strikingly, Calvo-Merino and colleagues reported the most
robust engagement of sensorimotor cortices, including dorsal and ventral premotor cortices, superior parietal lobe extending along the intraparietal sulcus, and the superior temporal sulcus, when ballet dancers watched ballet (but not capoeira) and when capoeira performers watched capoeira (but not ballet).

This dissociation of sensorimotor engagement based on longstanding expertise provided striking evidence in support of simulation theory, and inspired an innovative follow-up study by the same group of authors (Calvo-Merino et al. 2006). In this study, the authors sought to further examine the role of embodied expertise on perception, and to determine whether physical experience per se was necessary to shape perceivers’ brain responses when watching others in action (or whether it was perhaps possible that extensive visual experience alone could shape responses in a similar way). To address this question, Calvo-Merino and colleagues recruited professional male and female ballet dancers for an fMRI study, during which the dancers watched video clips of male and female gender-specific movements (i.e., movements that are only present in the repertoire of one gender), as well as movements that are performed by both male and female dancers. In order to determine the role of embodied expertise per se, the authors evaluated dancers’ brain activity when they watched movements that were specific to their gender compared to movements performed by the opposite gender (i.e., movements they had extensive visual familiarity with, but never performed). The authors found more robust engagement of premotor, parietal, and cerebellar cortices when dancers watched movements they had extensive familiarity performing (as well as seeing) compared to those movements that were every bit as visually familiar, but for which they were not motorically familiar (Calvo-Merino et al. 2006).

These two studies have sparked a number of follow-on studies that have continued to investigate how longstanding expertise shapes the brains and behaviour of perceivers, with subsequent work showing that trained dancers are better than non-dancers at discriminating pairs of point-light displays of dance movements (Calvo-Merino et al. 2010), and people can accurately recognize point-light displays of their friends versus strangers dancing (Loula et al. 2005). Together, this field of work demonstrates that longstanding expertise in the visual and especially the motor domain shapes how we perceive others moving around us.

### 22.2.2 De novo (laboratory) expertise

A useful counterpoint for examining how expertise changes how we see others’ actions is provided by studies that train participants to become experts with specific actions. As with investigations into longstanding expertise, in this domain as well researchers have enlisted the help of dancers and used dance learning paradigms to provide new insights on these questions. One of the first studies to do this was performed by my colleagues and myself, and involved tracking the brain activity of a company of professional contemporary dancers as they learned a complex new dance work (Cross et al. 2006). Across six weeks of the rehearsal period, we invited dancers into the laboratory to undergo an fMRI scan while they watched two types of choreography: a number of movements that they were currently in the middle of learning as part of their new choreography, and a group of kinematically similar “control” movements that the dancers never actually learned or rehearsed. Following each short video clip, the dancers were asked to rate how well they thought they could perform the movement they had just watched. Behaviourally, we found that the dancers unsurprisingly rated their own performance of the rehearsed (but not control) movements as improving across the course of the study. More strikingly, the brain imaging findings showed that as dancers watched movements they rated their ability to perform as increasingly better, clusters of activity within left ventral premotor and
inferior parietal lobule showed increasingly robust responses. My colleagues and I interpreted this as evidence that the better you are at performing an action that you have been learning in the recent past, the more you simulate that action. Here again, the simulation theory of action perception appears to be theoretically consistent with this finding.

Since this initial study, my colleagues and I performed a number of other studies involving laboratory-based dance training paradigms to address in more depth the relationship between quantifiable experience and how we perceive others in action. In one subsequent study, we examined how visuomotor and visual experience alone among novice dancers learning to play a dance video game shaped engagement of sensorimotor cortices. This study showed that both visual and visuomotor experience shaped parietal and premotor engagement in a similar manner (Cross et al. 2009). A more recent follow-up study also performed with novice dancers learning hip-hop dance sequences in the context of a popular video game probed this relationship in more depth, and found that the more sensory modalities through which one acquires experience with a new action, the more strongly sensorimotor cortices are engaged during action observation (Kirsch and Cross 2015). We have also examined how visual experience alone shapes perception of actions that are beyond observers’ motor abilities (such as mechanical movements made by wind-up toys, or complex gymnastic passes performed by expert rhythmic gymnasts), and report that visual training results in decreased recruitment of particularly visual regions associated with action observation (such as the extrastriate and fusiform body areas (Cross et al. 2013)). Even though such pre- and post-training fMRI experiments are costly in a number of ways, this research approach provides a useful counterpoint to longstanding expertise in building our understanding of how our experiences shape how we see others.

22.2.3 Expertise, embodiment, and aesthetics

One final research area that draws together action expertise and the performing arts that merits discussion here concerns how this relationship influences an observer’s aesthetic or affective responses. This research area continues to attract increasing research attention (Orgs et al. 2013, 2016; Kirsch et al. 2015, 2016a; Christensen et al. 2016a), not least because of various international initiatives that seek to bring together artists and scientists to generate truly interdisciplinary research from the ground up (such as the Dance Engaging Science initiative). Studies have shown that not only are sensorimotor cortices engaged when we observe dance movements we find pleasurable to watch (Calvo-Merino et al. 2008), but also that these brain regions can show increasingly robust responses the more we like watching a movement, and the less we can embody it (Cross et al. 2011).

Specifically, this latter study asked dance-naïve participants to watch a series of short video clips depicting ballet and contemporary dance movements, as performed by members of the Leipzig Ballet, and found that participants rated the movements they found most pleasing to watch as the most difficult to perform (a somewhat intuitive finding when one considers that patrons of the arts often pay substantial sums of money to watch expert artists perform physical feats they could never imagine doing themselves). According to simulation theory (and based on the studies reviewed in the previous sections), one might expect sensorimotor cortices to respond most robustly when observers watch things they can physically execute themselves; instead, we found robust engagement of the right intraparietal sulcus (as well as visual regions) when participants watched movements they rated as nearly impossible to perform (such as a male dancer performing a triple rotation jump or a female dancer performing a perfect split leap), but highly enjoyable to watch.
Subsequent work shows that dance expertise also leads to greater somatic manifestations of affect, as shown by implicit facial responses (Kirsch et al. 2016b), and a general measure of arousal revealed by galvanic skin responses (Christensen et al. 2016a). In addition, training studies document that when dance-naïve participants spend time coming to the laboratory each day and practising learning new dance sequences, their aesthetic preferences for these movements increase with increasing embodiment and ability to perform (Kirsch et al. 2013, 2015), and this relationship is present from early adolescence through to older age (Kirsch and Cross 2018). As discussed in a review article on this topic (Kirsch et al. 2016a), the relationship between an observer’s body, his or her motor skills, and how this shapes aesthetic preferences is only just beginning to be explored, and given the biological and evolutionary value of the human body, a more developed understanding of this relationship is likely to yield important insights into how we create and appreciate art as well.

### 22.3 Skill and expertise insights to optimize human–robot interactions

Another burgeoning area of inquiry greatly informed by research into skill and expertise is that of human–robot interaction. At first blush, this topic might seem remote from the theoretical underpinnings discussed in the first section of this chapter, and the insights gained from expert dancers in the second section. However, a growing number of laboratories are tackling challenges regarding designing socially engaging robots and how to optimize human–robot interaction by taking a social cognition-based approach that very much draws upon core ideas advanced by simulation theory. More specifically, a dominant view in social cognition states that throughout phylogeny and ontogeny, humans have developed to seek out self–other equivalences, which form the foundation of social cognition (Meltzoff and Prinz 2003; Meltzoff 2007). This account, known as the ‘like me’ hypothesis and consistent with simulationist accounts of action perception (cf. Gallese and Goldman 1998), further proposes that actions performed by oneself and another are represented in supramodal cognitive codes (Meltzoff 2007). Much of the literature covered in the first section of this chapter provides support for the ‘like me’ hypothesis, in that it demonstrates evidence of behavioural facilitation (Press et al. 2005; Catmur et al. 2007) and increased engagement of sensorimotor brain regions when individuals observe familiar actions or those with which they have expertise performing, or interact with agents similar to themselves (Buccino et al. 2004; Tai et al. 2004; Cross et al. 2006, 2009; Shimada 2010).

In addition to sensorimotor brain regions’ involvement in sharing experiences between actor and observer, successful interaction with others also entails taking an interaction partner’s perspective. A part of the brain known as the right temporoparietal junction (TPJ) is implicated in this process (Saxe and Kanwisher 2003; Apperly et al. 2004), and neuroimaging findings support the ‘like me’ nature of TPJ engagement when interacting with socially similar others (Klapper et al. 2014; Takahashi et al. 2014). Sensorimotor cortices along with TPJ are components of a broader network of cortical regions collectively called the ‘social brain’ (Frith 2007; Pelphrey and Carter 2008), and are involved in how we perceive, learn from, and interact with other agents we encounter in our social environments.

#### 22.3.1 From social cognition to social robotics

Over the past decade, individuals working to develop socially interactive robots are taking an increased interest in the social cognition and social neuroscience research reviewed above (Asada 2001, 2014; Breazeal 2007; Ishiguro 2013). An ongoing goal for robotics designers has been to...
maximize the similarity of artificial agents to humans, in terms of appearance and movement (while perhaps attempting to circumnavigate the uncanny valley; see Mori 2005), in an aim to make particular artificial agents as ‘like me’ as possible (Coradeschi et al. 2006). This idea also fits nicely with the notion that the expertise we have as being humans ourselves, and interacting with others like us, should be capitalized upon when creating artificial agents. In an elegant review paper, Press details many examples from behavioural psychology and cognitive neuroscience studies that support the notion that social brain regions are biologically tuned, demonstrating the most robust responses when viewing or interacting with other agents who look or move like us (Press 2011).

As compelling as the data and arguments supporting this idea that our experience of being human and interacting with other humans means that robots should be designed to be as human-like as possible are, it is important to note that a steadily growing body of research continues to call into question whether this relationship is really so straightforward. For example, my colleagues and I demonstrated that watching a human or robot perform actions that were robotic led to far more engagement of sensorimotor/social brain regions than smooth, familiar actions that participants had far more experience performing (Cross et al. 2012). These findings, as well as those reported by a number of other laboratories (Gazzola et al. 2007; Ramsey and Hamilton 2010) call for a far more systematic evaluation of the limits of neurocognitive plasticity when interacting with unfamiliar agents, in order to better understand how malleable social cognition is, the actual value of embodied expertise in shaping human–robot interactions, and the role that learning or longer-term exposure plays in shaping perception. As enthusiasm builds among the robotics community for incorporating social cognition and neuroscience research into robotics design (Hashimoto et al. 2006; Belpaeme et al. 2012; Lakatos et al. 2014; Henschel et al. 2020), future research examining the role of expertise should benefit the development of socially engaging and compatible robots by more clearly establishing the extent to which an artificial agent must be perceived as ‘like me’ in order to elicit a social response in a human interaction partner’s brain or behaviour (Hortensius and Cross 2018; Cross et al. 2019).

22.4 Conclusions

The aim of this chapter was to provide an overview of how embodying expertise as a performer or observer shapes our perception of others, and how research exploring this relationship is helping to spur developments in both arts and aesthetics as well as development of robots designed to socially engage with people. As stated previously, this chapter did not attempt to exhaustively cover all issues related to these domains, but instead aimed at providing the reader with a taste of the theoretical history and underpinnings for these questions, and some of the research approaches and findings that drive this field forward. Naturally, many exciting opportunities exist for follow-up work in both the performing arts and robotics domains. For example, future work in the arts domain could examine how the relationship between physical or embodied expertise and aesthetic preferences changes across time, either with the accumulation of additional skill, or with repeated exposure to the same stimuli over a longer time course than the studies described here (for example, many weeks or months rather than several days). It seems likely that the benefits that embodied expertise afford aesthetic preferences will only last up until a certain point, after which observers will become bored, no matter how much skill they have to perform the action they are observing. In the robotics domain, here as well great scope exists for the type of training studies described in the dance section, examining how the accumulation of social experience or interactive expertise with robots changes how people perceive these machines. Moreover, it will be illuminating to examine the extent to which
social similarity or expertise influences the utility of robots used in social contexts. In sum, both areas are ripe for future exploration, and a clear understanding of the theoretical and empirical foundations of embodied expertise will bolster progress in these domains.

Note

1 Capoeira is an Afro-Brazilian martial art that draws on elements of dance, self-defence, music, and acrobatics.

References


Expertise as a performer and perceiver


