Word Counts: Abstract – 229; Main Text – 14049; References – 3363; Entire Text – 17808

Tradition and Invention: The Bifocal Stance Theory of Cultural Evolution

Robert Jagiello[a] robert.jagiello@anthro.ox.ac.uk

Cecilia Heyes[b] cecilia.heyes@all-souls.ox.ac.uk

Harvey Whitehouse[a] harvey.whitehouse@anthro.ox.ac.uk

[a]Centre for the Study of Social Cohesion, School of Anthropology and Museum Ethnography, University of Oxford, Oxford, OX2 6PE, UK

[b]All Souls College and Department of Experimental Psychology, University of Oxford, Oxford OX1 4AL, UK

Short Abstract

Studies of cultural learning have tended to prioritize instrumental learning and innovation over conformism and the preservation of conventions. Here we argue for the Bifocal Stance Theory of cultural evolution which emphasizes both an instrumental stance in social learning, focused on the accomplishment of end goals via knowable (or potentially knowable) causal pathways, but also a ritual stance, focused on affiliating with other members of the group through the adoption of shared normative conventions. Systematic analysis of the social cues prompting these stances, allows us to present a new roadmap for future research into the cognitive foundations of cultural evolution.

Long Abstract

Cultural evolution depends on both innovation (the creation of new cultural variants by accident or design) and high-fidelity transmission (which preserves our accumulated knowledge and allows the storage of normative conventions). What is required is an overarching theory encompassing both dimensions, specifying the psychological motivations

and mechanisms involved. The Bifocal Stance Theory (BST) of cultural evolution proposes that the co-existence of innovative change and stable tradition results from our ability to adopt different motivational stances flexibly during social learning and transmission. We argue that the ways in which instrumental and ritual stances are adopted in cultural transmission, influence the nature and degree of copying fidelity and thus also patterns of cultural spread and stability at a population level over time. BST creates a unifying framework for interpreting the findings of otherwise seemingly disparate areas of inquiry, including social learning, cumulative culture, overimitation and ritual performance. We discuss the implications of BST for competing by-product accounts which assume that faithful copying is merely a side-effect of instrumental learning and action parsing. We also set out a novel 'cultural action framework' bringing to light aspects of social learning that have been relatively neglected by behavioural ecologists and evolutionary psychologists and establishing a roadmap for future research on this topic. The BST framework sheds new light on the cognitive underpinnings of cumulative cultural change, selection, and spread within an encompassing evolutionary framework.

Keywords: cumulative culture, cultural evolution, innovation, instrumental stance, ritual, ritual stance, social learning, tradition

1. Introduction

Scholars have long marvelled at the human capacity for creative invention, enabling our species to conquer new habitats and solve increasingly challenging problems through cumulative cultural evolution (Cavalli-Sforza & Feldman, 1981; Mesoudi & Thornton, 2018; Richerson & Boyd, 2005; Tennie, Call, & Tomasello, 2009). The effectiveness of tools and technologies often depends on the ability to respond flexibly to changing conditions, to *innovate* during the process of transmitting information, thus resulting in sequential improvements of tools, acts and artefacts—a phenomenon that has been researched extensively (e.g. Caldwell, Renner, & Atkinson, 2018). This emphasis on technological innovation, however, has tended to pay less attention to the evolutionary consequentiality of *non-instrumental* learning in the transmission of culture. For most of human history,

cultural evolution has not only been innovative but also characterized by slavish reproduction of rituals, traditions and conventions, resulting in the remarkable longevity of some of these cultural traits in the form of *cultural traditions* (Box 1). High-fidelity preservation of non-instrumental culture is just as much of a hallmark feature of our species as is technological invention through sequential change (Whitehouse, 2021). But what are the mechanisms that cause some traits to change within a few generations while others retain their form and stability for millennia? What are the respective functions of tradition and innovation in the evolution of culture and what are the mechanisms through which these are expressed?

Past research has shown the adaptiveness of certain biases in social learning, most notably regarding social cues such as skill, age and similarity to the learner (Mesoudi, 2011). For instance, the tendency to preferentially copy actions from individuals who display markers of success typically leads to improved propagation of optimally efficient technologies like arrowheads that maximize caloric returns during hunting (Atkisson, O'Brien, & Mesoudi, 2012). Such research is typically concerned with questions about *whom* and *what* are copied during the process of acquiring knowledge socially but seldom addresses factors that impact the fidelity of intergenerational transmission, or the extent to which the content changes as a result of being propagated. Here we build on previous research distinguishing ritual and instrumental stances in cultural evolution, which we refer to as 'Bifocal Stance Theory' (BST). During learning, ritual and instrumental stances are like bifocal spectacles, where "the upper half of the lens is used naturally when gazing at more distant objects (one might think of this as analogous to an instrumental perspective focusing on the bigger picture, oriented to end goals), whereas the lower part of the lens is better for examining things in close up (analogous to the ritual perspective focusing on detailed action parsing, oriented to the gestural level)" (Whitehouse, 2021: 35). Such a bifocal arrangement allows us to shift the focus back and forth when attending to different aspects of the learning environment.

Previous research on the BST framework has produced evidence that distinctive social cues, as well as certain features of action representation (e.g. the salience of end goals) trigger different motivational states in the minds of cultural learners, in turn influencing copying fidelity and rates of innovation (for early formulations of the BST framework drawing on

anthropological and developmental research, see Whitehouse, 2011, 2012b; for subsequent theoretical overviews, see Legare & Nielsen, 2015; Legare, 2019; Legare & Nielsen, 2020; Watson-Jones & Legare, 2016; Whitehouse, 2021: Chapter 1). BST prompts many questions with far reaching implications for cultural evolution theory that remain unaddressed, however. Here we seek to systematically discuss the cognitive underpinnings of social learning and the distinctive motivations that drive different modes of cultural transmission. We also consider underlying cognitive architecture, such as the level of deliberateness and domain-specificity with which individuals preferentially copy. This allows us to explore viable alternatives to popular assumptions in the current literature on convention learning – including, for instance, the possibility that our psychological capacity for high-fidelity ritual transmission is a culturally inherited "gadget" (Heyes, 2018a; 2019)—contrasting with past research that entertains innate learner motivations (e.g. Whitehouse, 2021) or even a Chomskean "universal ritual grammar" (Legare & Nielsen, 2020). Moreover, our approach sheds light on magical practices, which combine conventional actions with high goal salience. We also place BST into a wider research context by outlining its use and, clarifying its relevance to competing by-product accounts of overimitation and action parsing. We present a novel cultural action framework, which allows us to generate a series of testable predictions for consideration in future research, potentially bridging various sub-fields of cultural evolution research which, in the past, have tended to overemphasise either technological learning or normative conformism rather than both together. The extent to which BST is able to stimulate collaborative efforts between the disciplines lies in its ability to bring a more cognitively focused treatment of social learning into the ambit of anthropological and evolutionary theorizing. Conversely, it is also bringing ideas grounded in anthropology and evolutionary sciences to bear on various fields of cognitive science, such as developmental psychology which has long been posing pertinent questions about the different functions of social learning and imitation (Keupp, Behne, & Rakoczy, 2013; Schmidt, Butler, Heinz, & Tomasello, 2016; Uzgiris, 1981). We argue that transdisciplinary integration is essential to appreciate the flexibility with which humans alternate between varying degrees of high-fidelity copying and innovation, in order to solve both physical and social challenges—the main hurdles in the evolution of sociocultural complexity.

2. Bifocal Stance Theory (BST)

BST proposes a distinction between ritual and instrumental stances on observed behaviour which tap into distinct motivational systems, sensitive to different social cues, giving rise to different patterns of cultural transmission (Whitehouse, 2011, 2021). Adopting an instrumental stance, we expect observed behaviour to be directed towards the accomplishment of specifiable end-goals via potentially knowable causal pathways. In a ritual stance, we do not necessarily expect observed behaviour to be directed towards the accomplishment of specifiable end goals and, even if we do, we assume (implicitly or explicitly) that at least some of the actions required to perform the behaviour correctly do not contribute to those end-goals in ways that are potentially knowable. In other words, the instrumentality or conventionality of an action often lies in the eye of the beholder and the more the salience of causal pathways to an end goal is degraded, the more the observer's behaviour becomes oriented to social and affiliative outcomes, rather than to physical and technical ones. Participation in collective rituals signals membership of a cultural group and willingness to participate is at least partly motivated by the rewards of belonging.

Examples of behaviour that strongly activate the ritual stance range from relatively elaborate ceremonies such as weddings, funerals, baptisms, initiations, puberty rites, coronations, carnivals, and liturgies through to simpler behavioural conventions such as greetings, dining etiquette, codes of conduct, proper attire, and ceremonial precedence. All these forms of conventional behaviour comprise innumerable action clusters that lack specifiable end-goals or, even if they have goals, stipulate the observance of behavioural scripts that lack a fully knowable causal structure. The *puja* rituals performed by Jains in India, which involve various worshipful actions directed towards idols, illustrate this quite well (Humphrey & Laidlaw, 1994; Whitehouse, 2004: 95-97). Different variants of the *puja* range from bathing an idol (*prakshal puja*) to placing a flower on it (*pushpa puja*). When asking different practitioners about why they engage in this practice, Humphrey and Laidlaw documented a multitude of different meanings, ranging from symbolic and expressive motivations (e.g. to enable one's spiritual knowledge to blossom or become purified like a flower), to more decidedly instrumental interpretations (such as that the scent of the flower makes the process of worship more pleasant). Interestingly however, some informants

maintained that the placing of flowers did not mean anything at all and needed to be carried out simply because that is what Jain worshippers do (Humphrey & Laidlaw, 1994: 35).

Anthropologists have long debated what makes ritual behaviour distinct from any other and have focused on variants of the idea that rituals are governed by "rules without meaning" (Staal, 1989) or that intentional meaning is not intrinsic to the actions performed (Humphrey & Laidlaw, 1994). They note that certain procedures must be carried out in stipulated ways (e.g. the bride and groom must dress a certain way, the person to be knighted must kneel in order to be tapped on the shoulder by a sword, and so on) and the reasons for those actions cannot be derived from the intentional states of the actor and seem instead to come from some other source (tradition, the ancestors, a divinity, etc.). BST proposes that a rationale for following the ritual script is inessential for successful and high-fidelity ritual transmission and may be entirely lacking in some cultural traditions. For instance, when a Catholic first learns to cross herself upon entering a church, this likely constitutes the type of automated and unquestioned behaviour that precedes fully established and explicit knowledge of the wider religious belief system the act is embedded in (Whitehouse, 2002). In other words, even though the congregant may or may not acquire a goal-oriented rationale later on, the initial instances of behaviour acquisition are imitative and driven by the motivation of simply wanting to fit in with the other church goers. Thus, ritual participants may carry out the required procedures without at first being able to say why. They may not know what the actions 'mean' and may not see them as linked to any kind of specifiable origin or future end-goal.

Even where end-goals of ritual and tradition are made salient early on during learning, they do not explain 'recurrent fidelity' (Heyes, 2018), the longevity of these practices from an evolutionary point of view. In a hypothetical world where only the instrumental stance exists (where cultural transmission is goal-oriented at all times), selection should favour accurate copying to prevent technologies from regressing to more primitive states but also goal emulation, which eliminates the steps of a sequence that are inconsequential in reaching a goal. Over time, this should have led to the gradual disappearance of practices that have a weak action-outcome link (practices that do not surpass chance levels in achieving a desired end state). BST proposes that the ongoing and stable presence of practices that are ineffective in producing environmental outcomes, result from a non-

instrumental, affiliative motivation, helping to regulate social life in groups (such as hierarchy maintenance or improved coordination). Viewed in this light, participating in a collective ritual is very different psychologically from the experience of working together on an instrumental task, such as fixing a vehicle or building a fence.

It is important to emphasize, however, that BST does not attempt to categorize socially learned behaviour into either ritual or instrumental actions. We argue, on the contrary, that most social learning involves a complex mixture of both ritual and instrumental stances, activated flexibly, often in rapid succession. Action sequences may exhibit properties that make them more likely to activate either a ritual or instrumental stance, e.g. if the start and end points of a sequence turn out to be the same, learners are more inclined to interpret the action as *causally opaque* (Box 1.) and thus conventional (Watson-Jones et al. 2016). But whether such start-end state equivalence activates a ritual stance also likely depends on the presence of other contextual cues.

Box 1 summarises the key concepts required to grasp this framework. When observing behaviour through the lens of the ritual stance, the focus of attention is not on acquiring a causal rationale for the procedures. Social etiquette, clothing fashions, tea ceremonies, and even the rules of childhood games *may* be ascribed a purpose (whether in the process of teaching and learning or in reflecting later on why our habits take the form that they do) but often their purposes remain mysterious, and the actions are simply copied without question. Whether or not rituals are attributed purpose, it would make little sense to try to formulate in rational causal terms how that purpose is realized (Whitehouse, 2011, 2012b, 2021). Interpreting an action sequence through the lens of the instrumental stance on the other hand (either via cues provided by the model or properties of the action itself) prompts an expectation of learning something technically useful, thus directing attention to the causal structure of behaviour and its end-point.

As we will discuss below (in section 2.3), some cultural practices are likely to activate both ritual and instrumental stances. These are goal-directed rituals, ranging from magical spells to sporting contests. They incorporate cues that signal instrumentality (a salient goal) as well as cues that signal conventionality (e.g. irresolvable opacity), which make the adoption

of either stance more likely than in other cases that can be viewed as more decidedly conventional or technological.

Box 1. Key terms and their definitions

Convention, Tradition: Actions that regulate and maintain social life in groups through various affiliative functions, e.g. by serving as group identity markers or coordination devices. Such actions are typically reproduced over long periods of time through faithful intergenerational transmission.
Innovation: Actions that deviate from modelled forms either through transmission error or purposeful modification aimed at achieving goals in better ways.
Quasi-instrumental Rituals: Actions that signal conventionality but also promise the attainment of a particular end-state.
Social Learning: Information acquisition through interaction with- and observation of other individuals and their products (Heyes, 2012)
Social Learning Bias: A learning strategy that prioritizes some types of information over others (content biases) or which leads to preferential learning from models based on their attributes, such as displayed competence (model biases; Mesoudi, 2008; Atkisson, O'Brien, & Mesoudi, 2012;)
Overimitation: Imitation of causally irrelevant steps in an action sequence (Lyons, Young, & Keil, 2007; Over & Carpenter, 2009)
Causal Opacity: An action sequence is causally opaque when causal relations among its components are not apparent to an observer (e.g. when a step is seemingly physically inconsequential to the one that follows)(Whitehouse, 2011; Legare et al. 2015; Kapitány& Nielsen, 2019)
Goal Demotion: Absence of a salient goal in an observed action sequence (Nielsen, Tomaselli, & Kapitány, 2018)
Ritual Stance: A detail-focussed mode of processing observed behaviour activated by cues indicating that the action is irresolvably causally opaque and/or goal-demoted. The ritual stance promotes social learning that has an affiliative function(Whitehouse, 2012b). It has also been described as the 'normative stance' (Whitehouse, 2011, 2012b)
Instrumental Stance: A goal-oriented mode of processing observed behaviour activated by cues indicating that each step contributes in a knowable way to the outcome (see Opacity Resolvability). The instrumental stance promotes social learning of technical skills (Whitehouse, 2011).
Bifocal Stance Theory (BST): The framework in which cultural transmission is guided by both ritual and instrumental stances
Opacity Resolvability/Causal Knowability: The distinction between actions that are causally opaque and those that only appear to be causally opaque due to a lack of physical knowledge (resolvable opacity) (Whitehouse, 2011, 2012b). Thus the causal structure of an action sequence can be viewed as either knowable or not.
Domain-specificity: A cognitive mechanism is domain-specific when it is specialized for the performance of distinctive tasks (Cosmides & Tooby, 1994)
Domain-generality: A cognitive mechanism is domain-general when it uses the same computations to perform a range of different tasks
Cognitive Gadget: A distinctively human cognitive mechanism specialised predominantly by cultural rather than genetic evolution (Heyes, 2018)

Molecular view: A fine-grained, detail-focussed way of observing behaviour, sensitive to its constituent parts (e.g. the discrete steps that make up the action sequence)
Molar view: A coarse-grained, goal-orientated way of observing behaviour, sensitive to the action sequence as a whole and its endpoint.

2.1 Stance selection

Any observed behaviour that is not obviously contributing in a causally transparent way to a readily identifiable end-goal may be interpreted through the lens of either the ritual or the instrumental stance, guided by salient social cues, such as comportment, gait, gesture, ostensive cueing, social synchrony, coordination, confidence, experience, seniority, or verbal cues suggesting normativity or expectations of conformism. To illustrate how difficult it can be to select a ritual or instrumental stance on observed behaviour, it is instructive to consider the example of 'Sylvia's recipe' (Gergely & Csibra, 2006; see also Whitehouse 2011 and 2021). Sylvia had a distinctive way of roasting ham. She cut off both ends of the joint before placing it in the oven. She had learned this technique at her mother's knee, never questioning it. Many years later she prepared this dish while her elderly mother was visiting. Observing Sylvia's technique in astonishment, the mother asked her what she was doing. It turned out that when Sylvia was a child the family roasting tin was too small to accommodate an average joint of ham and that is why her mother cut off the ends. The young Sylvia observing her mother's cooking methods never enquired as to the underlying rationale. She may have assumed that the removal of the ends of the joint had some instrumental explanation (e.g. to allow the juices of the meat to flow out) even if that purpose was not immediately obvious. One could equally imagine a cultural milieu in which Sylvia's recipe was accorded a supernatural function, for example to release the spirit of the animal that has given its flesh (quasi-instrumental – see 2.4 below). But it is equally possible that Sylvia adopted a purely normative, non-instrumental understanding of the procedure, prioritizing its *social* meanings and functions. Accordingly, cutting off the ends could be a clue to the cook's ethnic origins, via traditions passed down by many generations of mothers before her. Moreover, whenever a normative perspective is activated there is also a further, if rather baffling, possibility: that nobody knows why meat should be prepared in this peculiar fashion, it simply *should*. Such a response to stipulated ritual procedures is

particularly common when the procedures themselves have become familiar habits. In many such cases, scripts that are so familiar that they become sedimented in implicit procedural memory no longer require conscious effort to perform and are thus less likely to prompt conscious reflection on why we do them at all (Whitehouse, 2004). When we interpret behaviour in these ways—as a normative convention rather than as an instrumental technique—we adopt a ritual stance (Whitehouse, 2012b;McKay & Whitehouse, 2015).

The story of Sylvia's recipe is instructive because it shows clearly that many observed behaviours could activate either the ritual stance or an instrumental stance – social learning could potentially go either way. It might well be the case, that knowing *instrumentally* how best to release liquids from meat during the roasting process with the intention of maximizing the tenderness of the resulting dish, could confer prestige on the cook. Clearly, any prestige acquired in that way is a consequence of demonstrated competence at cooking rather than cooperative or affiliative commitment. Nevertheless, cutting off the ends of the meat could just as credibly be understood as a ritual action – that is, a convention that cannot be justified because it leads to an outcome via some knowable physical-causal chain. This would be the case if, for example, Sylvia's mother is thought to have removed the ends of the meat to make it recognizable as a dish associated with a particular cultural group. In some cases, the 'done' or 'proper' way of presenting food is also associated with an elite stratum in society (e.g. the serving of chilled oysters on the shell) and, since this is the 'high class' or 'posh' way of doing it, dutifully observing this aspect of the ritualized presentation of food may indeed confer prestige both on the person who has mastered the method and the discerning consumer of the dish.

BST proposes that when we learn or carry out a particular sequence of actions we "flexibly switch between the two based on relevant social cues" (Hermann, Legare, Harris, & Whitehouse, 2013). But even though an action sequence may entail the activation of both ritual and instrumental stances at different moments, one or other of the two stances may tend to predominate in a particular action cluster, such that we are more inclined to view it through the lens of the ritual stance (e.g. when kneeling to pray in church, in conformity with the congregation) or the instrumental stance (e.g. when kneeling to fix a bicycle, to

facilitate access to a low-lying sprocket). Nevertheless, there are also many action clusters that activate both stances in such rapid succession that it may be hard to decide whether to class them overall as ritual or instrumental. The noted anthropologist Edmund Leach, for example, describes the process of gardening among the Kachin of highland Burma as continually oscillating between practical instrumental elements necessary to ensure plant growth and 'aesthetic frills' that are technically unnecessary but conventional aspects of horticulture, identifying the gardener as a *bona fide* member of the Kachin cultural group (Leach, 1954). Much the same may be said of spectator sports such as football (Newson et al., 2020), which have many very ritualistic aspects (from 'Mexican waves' to the wearing of special scarves) but also very causally transparent procedures for achieving end-goals (such as the placement of a ball at the back of a net). We return briefly to this point below when discussing the affiliative aspects of ritual participation and also in connection with quasi-instrumental rituals.

The use of a 'stance' metaphor for ritual and instrumental social learning differs from Dennett's (1987) notion of an 'intentional stance', which suggests a mindset cognisant of the view that the actions of others are governed by beliefs and desires. As discussed in the next section, we do not assume that ritual and instrumental stances operate only in ways that are deliberative and accessible to conscious awareness. We propose instead that the stances are best understood as states of increased attentiveness towards either the link between conventionality cues (Figure 1) and positive social outcome or the link between instrumentality cues (Figure 1) and improved goal achievement. In other words, the learner's sensitivity to properties of actions and the models performing them, and propensity to associate these properties with either conventionality or instrumentality, triggers differential learning approaches. In the case of triggered conventionality, the learner is more attentive to action details (such as the chronological order of the action sequence) resulting in overall higher copying fidelity and greater longevity of the modelled action during transmission, while in a scenario where the action is processed as instrumental, attention shifts more towards the outcome, thus resulting in overall lower copying fidelity and higher rates of innovation.

Importantly, this does not mean that the instrumental stance is incapable of producing stable practices. On the contrary, techniques that have proved reliable in achieving a certain

goal often become fixed, preventing regression to more primitive states (Tennie et al., 2009). Conversely, innovation within the ritual domain occurs as well, and deviations from a script may even be encouraged, for example in the case of Congolese spirit play (Lewis, 2002) or status competition among the Tiwi (Hart, Pilling, & Goodale, 1988). Indeed, the ritual and instrumental stances are intimately interwoven in many human pursuits, we anticipate that both forms of social learning overlap and influence each other. For example, when learning a particularly complex technical skill, it helps to mix in a little bit of ritual. While too much ritualization impairs learning, perhaps due to cognitive load, a small amount of ritualization may enhance it (Kapitány, Kavanagh, Whitehouse, & Nielsen, 2018). Given examples of ritual innovation, technological stability, and stance admixture, the differences in transmission fidelity between the two stances are predominantly one of degree. All else being equal, practices that inspire copying via the ritual stance tend to be more accurately reproduced *relative* to those copied via the instrumental stance, making the former comparatively more resistant to change.

Overall, BST therefore seeks to explain the co-existence of adherence to convention and technological innovation in humans from an evolutionary perspective, where certain cues of the copied action influence the nature and degree of copying fidelity.

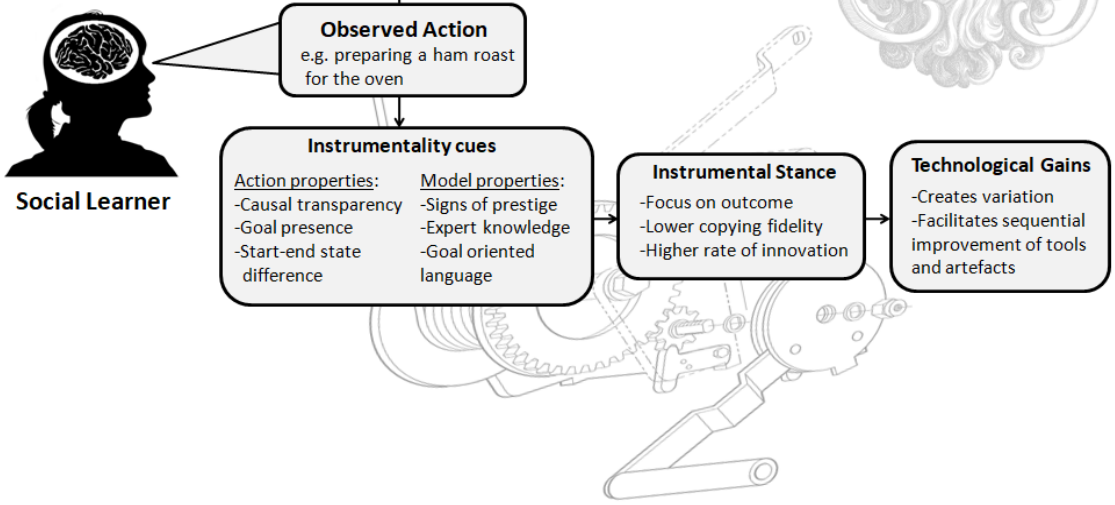

Figure 1. Schematic overview of the Bifocal Stance Theory (BST)

2.2 Stance deliberateness

A fundamental question prompted by the BST framework, largely unaddressed in previous literature (although see Whitehouse, 2002), is the degree of deliberateness with which instrumental and affiliative motivations operate. The stances could be automatic in the sense that they rely on psychological processes to which learners have little conscious access, or deliberative in that they rely on conscious cost-benefit calculations. More specifically, the stances could rely on Type 1 or Type 2 psychological processes (Evans & Stanovich, 2013; Kahneman, 2003; Norman & Shallice, 1986). Type 1 psychological processes do not depend on working memory; they are typically automatic, fast, associative, effortless and non-conscious. Type 2 processes make demands on working memory; they are typically deliberative, slow, rule-based, more effortful, and conscious.

It is easy to imagine the stances running on deliberative Type 2 processes. In this case, a boy watching an adult cross herself in church would make conscious decisions about the purpose and resolvability of the crossing action. Using explicit mentalising (also known as 'mindreading' and 'theory of mind'), he asks himself what the adult intends to achieve through her action, and what she wants to communicate to him (Csibra & Gergely, 2009). Using causal reasoning, he asks whether the action generates an instrumental outcome and, if so, mentalising again, whether anyone understands the action-outcome relationship. He may even engage in normative reasoning, asking whether the crossing action is required, encouraged, or prohibited for him by his social group. Based on these decisions – outputs of Type 2 mentalising, causal reasoning and normative reasoning processes – the boy adopts a stance, attending to the details (ritual) or overall form (instrumental) of the action, and copies it with greater or lesser fidelity with a specific motivation. For example, he may copy the crossing action with high fidelity expecting it to bring him divine grace, acceptance by his community, or the approval of the adult he is copying.

It is harder to imagine the stances running on automatic Type 1 processes because these processes are not part of our intuitive psychology. They have been discovered by cognitive science rather than revealed by common sense. In the automatic case, a stance is triggered rather than chosen. Observable features of the action and context, the conventionality and instrumentality cues listed in Figure 1, send the cognitive system into one of the two stances via associative processes. For example, when the boy sees the adult performing an action where the start and end states are equivalent- her hands are joined in prayer when she begins to cross herself and come to rest in the same position - it activates the cluster of psychological processes that constitute the ritual stance, including attention to detail and a diffuse expectation of social rather than non-social rewards. He acts expecting to get something good from other people but does not know what it will be or have a theory about why he will get it. Like the associative links that mediate faster responding with the left hand to stimuli on the left of the visual field (spatial compatibility effects; Lu & Proctor, 1995) and faster reading of colour words printed in the colours they name (Stroop effects, MacLeod, 1991) the associative links that mediate automatic activation of stances could be innate or learned. More precisely, they could be genetically or culturally inherited (see Section 5).

Everyday experience suggests that stance switching is at least sometimes deliberative in older children and adults. Most of us can remember asking ourselves in a new social context how faithfully to copy the locals. As a guest at a formal college dinner, should I wear an academic gown or just dress smartly; should I dip my fingers in the finger bowl, or just moisten my napkin? However, automatic stance switching would have significant advantages because it is faster and less cognitively demanding than deliberation. It would allow adaptive modulation of social learning in infancy and early childhood, before Type 2 executive processes have matured, and agile movement between stances throughout the lifetime. Therefore, it is likely that stance switching, like linguistic "code-switching", is sometimes deliberative and sometimes automatic. In the case of code-switching, bilingual friends move back and forth between their languages, within a conversation or even within a sentence. Switching can be deliberate - for example, a speaker may decide consciously to switch to English for scientific terms which she knows are more familiar to the listener in that language – but switches can also be triggered automatically by the emotional tone of the conversation, or a quizzical look from the listener (Pietikäinen, 2014).

Priorities for future research are to discover whether stances can be selected automatically and, if so, when stance switching is automatic and when it is deliberative. To find out, we need to ask people why they have copied with high or low fidelity, and to check whether their answers match the cue conditions that modulated their copying behaviour. A good match would suggest deliberation, and a poor match would suggest that stance selection occurred automatically, and their statements were rationalisations after the fact (see section 2.4.3 on the difference between attributed and evolved functions). Similarly, children and adults could be given tasks that demand working memory alongside social learning tasks that call for stance switching. If these concurrent tasks do not interfere with efficient stance switching, it would suggest automaticity. If efficiency declines as working memory demand increases, it would suggest that, in the tested population and circumstances, stance switching is deliberative. Based on studies of mentalising (Apperly, 2010; Heyes & Frith, 2014), metacognition (Goupil & Kouider, 2019; Heyes, Bang, Shea, Frith, & Fleming, 2020), and other cognitive processes that come in automatic (implicit) and deliberative (explicit) forms, it is likely that stance switching is automatic in infancy and early

childhood and then becomes more deliberative, but only in circumstances where there is minimal time pressure and significant risk that automatic processes will produce error.

2.3 The ritual stance, affiliation, and group identity

Why would the copying of irresolvably opaque behaviour be motivated by the desire to affiliate? Rituals serve many social functions, as exemplified by scarification procedures that signal mate quality (Singh & Bronstad, 1997), ceremonies that commemorate and promote acts of heroism (Kertzer, 1988), as well as resource distribution practices that buffer against inequalities due to exogenous factors (Woodburn, 1982). Despite the multitude of different purposes rituals can serve, their common underlying thread is the regulation of group life—it is through affiliation that the individual gains access to the group as well as the benefits that come with it (resources, mates, cooperation, protection, etc.) (Durkheim, 1965; Kertzer, 1989; Khaldūn, 1958). We propose several reasons why ritual participation leads to group bonding. One is that arbitrary conventions serve as exclusive identity markers for members of the ingroup. Although (or indeed *because*) such behaviours may have precisely zero value from a technical-instrumental perspective, participation is an attractive prospect for anyone wishing to affiliate with the ritual community in question. This explains how rituals survive the selective environment of our cultural landscape; they are not replaceable by actions construed in instrumental terms because they fulfil an entirely different function, namely that of maintaining patterns of cooperation, ranging from group loyalty to obligations of reciprocity, conferring a selective advantage on the cultural groups adopting them (although we also discuss the case of magical procedures below). As such, rituals establish *discrete ingroups*, in contrast with useful technologies that spread willy-nilly *across* groups (Whitehouse, 2012a).

Rituals also commonly generate social cohesion and thus cooperation via social synchrony, in the form of dancing in rhythm, choral singing, chanting, swaying, and marching, evincing well-documented effects on social bonding and cooperation (Catmur & Heyes, 2013; Fessler & Holbrook, 2016; Hagen & Bryant, 2003; Hove & Risen, 2009; Mehr, Krasnow, Bryant, & Hagen, 2020; Wiltermuth & Heath, 2009). Rituals typically require significant investment of labour and materials despite lacking any technical value or output and, as such, serve as

costly signals of commitment to the group (Irons, 2001; Sosis & Alcorta, 2003). Indeed, the more costly and fitness-decreasing rituals become, the harder it is to square participation with dislike for the group, enabling cognitive dissonance effects to motivate group love (Aronson & Mills, 1959). Moreover, modulating the frequency and emotional intensity of collective rituals also produces distinctive effects on the intensity and scale of group bonding: high-frequency, low-arousal rituals facilitate rapid spread of stable identity markers to large populations (such as doctrinal religions and nations) while low-frequency, high-arousal rituals generate highly cohesive localized groups (such as initiation cults, military brigades, and terrorist cells) (Whitehouse, 1995, 2000, 2004, 2018, 2021).

A growing body of empirical research suggests that the ritual stance is sensitive to social cues indicating that modelled behaviour is conventional, normative, and otherwise relevant to group alignment, and therefore that the motivation to copy such behaviour is affiliative in orientation (Clegg & Legare, 2016b; Herrmann, Legare, Harris, & Whitehouse, 2013). We therefore expect role models who are strongly associated with the group, for example because they are popular or exemplify group values, to trigger the ritual stance among would-be learners. By contrast, the instrumental stance is sensitive to social cues indicating that the behaviour is technically useful and copying is rewarded by the acquisition of new practical skills. Thus, we would expect role models who signal competence, experience, and skill to trigger the instrumental stance among would-be learners.

2.4. BST and quasi-instrumental rituals

Practices signalling conventionality but also promising to deliver desirable end-states are frequently reported in the ethnographic record. For example, competitive sports are oriented to outcomes, such as scoring goals, but in conventional, causally opaque ways (e.g. playing a ball into the opposing team's net but without using one's hands). Quasi-instrumentality is also a feature of conventional solutions to coordination problems (e.g. you must only drive on the left to avoid crashing into other vehicles). But it applies also to magical rituals (e.g. you have to utter a particular spell in order to make it rain). In all such cases, part of the causal pathway from action to intended outcome is underspecified (causally opaque). Most importantly, nobody expects there to be a causally transparent

reason why you use feet and not hands in one sport, drive on the left not the right in some countries, or repeat this incantation rather than that to make it rain. It is simply the correct way of doing things in this region or that cultural setting – the recognized convention. To the extent that such behaviours activate the ritual stance, they can certainly serve to communicate information about the group alignments of the person adopting them. They can tell us, for example, what sports people grew up with, what countries they live in, and what cultural traditions they uphold. The process of observing these conventions therefore has an affiliative aspect. Nevertheless, in all these cases the instrumental stance is also likely to be activated because there are very salient end-goals in play – such as communicating information, placing a ball in a net, driving safely, or ending a drought. Accordingly, this subsection discusses common examples from everyday life (such as games and the acquisition of language), but also focuses on how cases of magic can be accounted by BST, because so many rituals throughout the world's cultures (past and present) are concerned with bringing about end-goals.

2.4.1. Games and language

Boardgames and pretend play provide rich sources of cultural examples of goal-focused activities in which successful participation relies on adherence to previously established group conventions (Rakoczy, Warneken, & Tomasello, 2009). In chess, even though the explicitly stated goal is to 'checkmate' the opponent's king, the pieces can only be moved based on a collectively agreed set of arbitrary rules that do not stand in any resolvable physical-causal relation to that goal (e.g. there is no non-arbitrary reason why a bishop would only be able to move diagonally across the board other than the stipulated rule that this should be so). Most importantly, in much the same way as navigating traffic, learning chess requires high fidelity transmission of rules which are resistant to innovation—a player cannot suddenly decide that a piece moves differently than what is set out in the rules.

Analogously, pretend-play paradigms not only demonstrate children's propensity to adhere to non-instrumental rules but also their willingness to pick up specialized vocabulary that is exclusive to a discrete context. For example, Rakoczy, Warneken, and Tomasello (2008)

taught children to associate a specific action sequence with the term *daxing,* and later observed a puppet perform a different action while applying the same name to it. Across multiple iterations, children protest whenever these action-terminology incongruences occur, suggesting a ritual or normative stance in the context of the game but also in vocabulary acquisition more generally. Indeed, language acquisition entails similar requirements to game participation or the solving of coordination problems: learners must be willing to associate arbitrary but conventional and prescriptive phonemes with certain features of their environment—there is nothing inherent to a word that signals what it means. Much like the collective agreement that a chess piece can move in a certain way regulates the interaction between the two players, the language we acquire determines whom we can interact with, thus serving as an efficient identity marker and coordination device. Indeed, at least some aspects of children's word learning appears sensitive to the conventional nature of language (Graham, Stock, & Henderson, 2006) and young learners appreciate that newly acquired vocabulary can be understood by other members of their speech community (Diesendruck & Markson, 2001; Sabbagh & Henderson, 2007). Children also appear less inclined to learn a novel word when there is evidence that it is not shared by other speakers (Koenig, Cle, & Harris, 2004). Even where a language is less constraining by being spoken across different groups, local accents and dialects are often indicators of cultural origin. Past research has demonstrated an accent's ability to act as a hard to fake signal that is readily picked up by children, influencing their preferences for word utterances (Butler, Floccia, Goslin, Panneton, & Tech, 2011) and related objects (Mulak, Best, Tyler, Kitamura, & Irwin, 2014), as well as guiding their decisions about which models to trust and learn from (Kinzler, Corriveau, & Harris, 2011; Kinzler, Shutts, Dejesus, & Spelke, 2009). Given these affiliative motivations, it is plausible that language learning shares considerable overlap with the acquisition of cultural rituals and traditions.

2.4.2 Magic and Resolvability

In order to explain the acquisition of magical rituals through the stances, we turn to the role of action resolvability (Box 1). The *resolvability* of an observed action (Whitehouse,2011; Legare & Herrmann, 2013; Legare & Souza, 2014; Legare & Souza, 2012) concerns the extent

to which the causal structure of an action is perceived to be knowable (in theory if not de facto) and thus whether a novice will seek to resolve uncertainties by trying to establish how the learned behaviour 'works'. There are many goal-oriented actions in day-to-day life which are causally opaque to the observer but which are in fact resolvable (Whitehouse, 2011). For instance, unless one possesses basic mechanical knowledge, the causal link between moving the gear stick in a car and the effect it has on speed and acceleration is generally obscure. However, it is assumed by the observer (either implicitly or explicitly) that, in principle, there is information available that explains why gears must be shifted in a certain way in order to go faster. On the other hand, the ritual stance operates on the assumption that a sequence has no knowable physical-causal structure (i.e. its causal opacity is irresolvable), but is used within the social realm as a means to affiliate, bond, and to promote cohesion. Much like with perceived instrumentality and causal opacity, we argue that resolvability lies in the eye of the beholder—in principle a learner can perceive the causal relationship between flicking a switch and a light coming on (which is objectively resolvable, given sufficient understanding of how electrical circuits work) as *unresolvable* when relevant information about electricity is unavailable but contextual cues that favour a ritual stance are present (e.g. as may have occurred during first contact in highland New Guinea; Lawrence, 1989). Conversely, in the case of an action sequence typically regarded as irresolvably opaque, an agent may nevertheless interpret the sequence as resolvable if learning cues render an instrumental stance more likely – for example, if I am told that handshaking is not an arbitrary convention but a way of demonstrating that one is unarmed and thus unthreatening.

Magical rituals are abound in the ethnographic and historical record and may take a great diversity of forms, from the simplest rubbing boards of the African Azande used to detect witchcraft (Evans Pritchard,1937) to the highly elaborate seafaring magic of the Melanesian Trobriand islanders (Malinowski, 1935). What most magical rituals have in common, however, is the end-goal of warding off misfortune by supernatural means. The types of misfortune such rituals seek to prevent or reverse are immensely variable – from diseases of the body, mind, and failing relationships through to the deleterious effects of plague, pestilence, intergroup conflict, drought, fire, and flood. People everywhere recognise that many of these kinds of misfortune can be prevented by quite straightforward instrumental

efforts via natural causation such as medical treatments, disease prevention, fertilizers, insecticides, flood defences, irrigation systems, and so on. Nevertheless, such efforts to ward off misfortune are not always enough and when the risk of failure is great and the means of preventing it limited, people frequently turn to magic and the ritual stance. Hopes are pinned on causal pathways that cannot be fully specified and are thus irresolvably opaque. In other words, such magical thinking often appeals not so much to instrumental processes of cause and effect via mechanistic principles but to the petitioning of supernatural agents and the invocation of reciprocal obligations and other kinds of cooperative principles. There is also a stronger element of affiliative concern and social bonding in quasi-instrumental practices than in more purely instrumental techniques and interventions. Thus, the providers of irresolvably causally opaque medical treatments (homeopaths, shamans, witchdoctors, etc.) invariably foster closer and more affiliative relationships with their clientele than providers of purely instrumental cures.

Accordingly, BST proposes that, like Silvia's recipe, magical practices can recruit the ritual stance or the instrumental stance depending on the relative number of perceived cues that convey conventionality or instrumentality. For example, if a learner observes a magical practice with a clearly defined end-goal (instrumentality cue), which is nonetheless accompanied by normative language (conventionality cue), as well as performed by one or multiple highly similar models from one's own group (also a conventionality cue), then BST expects that this makes the activation of a ritual stance which prompts higher fidelity copying more likely. Similarly, an individual who is enlisting the services of a witch doctor, might view the practice through an instrumental stance if, despite the irresolvable opacity, there is a strong focus on the end-goal combined with the perception that the witch doctor is an expert (both cues to instrumentality). Rituals performed in the context of baseball may also contain goal-focused cues to such extent that they overshadow the irresolvable nature of the action, thus allowing them to be viewed instrumentally.

Accordingly, the differentiation between emic resolvability (the learner's perspective of whether an action is causally opaque in a resolvable manner) and etic resolvability (whether the actions sequence is actually resolvable) comes with an important implication for researchers: A practice which appears from the outsider's perspective to be magical (irresolvably opaque) may not be interpreted as such by the population that engages in it

(for example, because the instrumental stance is triggered during its transmission due to a variety of cues that make end-goals and assumed expertise of the model more salient). Consequently, it is important that researchers not only view ritual practices from an etic perspective, considering their evolved functions, but also from an emic perspective, taking into account the psychological processes of the learner. Moreover, in the next section we turn towards discussing how the evolved reason a magical practice is transmitted can differ substantially from the explanations given by its practitioners in the context of stance deliberateness.

2.4.3 Cultural selection and evolution

BST proposes that stances are adaptive - having two modes of social learning, for conventional and instrumental behaviour, makes individuals and societies better able to thrive in their physical and social environments. Stance psychology could have evolved via biological or cultural selection. If stance psychology is a "cognitive instinct" (Pinker, 2010), individuals genetically inherit a very specific propensity to develop the two modes of social learning, and those who inherit a better version have more biological offspring. If, however, stance psychology is a "cognitive gadget" (Heyes, 2018a; 2019), individuals genetically inherit only domain-general psychological resources, for example, attentional, learning and motivational processes that perform a variety of jobs. Using these domain-general processes, individuals learn stance psychology from older members of their community, and communities with a better version expand or proliferate (see Section 5). Whether stance psychology is shaped by genetic or cultural selection, deliberative or automatic, there is significant potential for the emic (attributed) functions of rituals to diverge from their etic (evolved) functions.

Although rituals may be performed for a consciously formulated reason (e.g. dancing *to make it rain*), we argue that implicit motivations (such as obtaining affiliative rewards) do not need to align with such attributed purposes. By associating the high fidelity copying of actions that bear conventionality cues (such as end-start equivalence) with positive social outcomes (such as being included by the group), automatic copying can guarantee that a practice is propagated with the level of fidelity that is required in order to fulfil its

etic/evolved functions, without the need of an emic/attributed function (which may or may not differ from the etic one). For example, singing in private may be consciously attributed a purely goal-focused (if casually opaque) function (e.g. to encourage flowers to grow), but we argue that the process by which this song was acquired would have been a social and normative one, whereby the learner must have shown willingness to copy arbitrary words and melodies accurately via the ritual stance. Therefore, a consciously formulated function may differ quite substantially from the learner's implicit motivations which are likely selected for by the practice's culturally evolved function. Thus, costly practices such as magic do not persist because of their efficiency in solving instrumental problems (it is unlikely a song will be transmitted across generations by virtue of its emically conceived function of making flowers grow), but we argue that, among other benefits, their longevity within a group can be due to their contributions to the practitioners' collective identities.

Cultural selection is often guided not by individual goals and preferences but by environmental pressures. For example, low-frequency and high-arousal practices (imagistic rituals) may not have been intentionally designed for the purpose of maintaining cohesion in small groups, but may rather result from selective forces of cultural winnowing such as inter-group conflict or high-risk subsistence strategies that favoured those types of rituals over others (Whitehouse, 2018). Accordingly, an automatic account of BST (as discussed in section 2.2) works very well in explaining how goal-oriented practices that are irresolvably opaque can still be copied for affiliative reasons, simply because copying similarly opaque practices with high fidelity has yielded social rewards in the learner's past. This is in line with research that demonstrates equal levels of affiliative mimicry across conditions where a goal is consciously known or not (Lakin & Chartrand, 2003). In other words, affiliative copying does not have to rely on deliberate goal focus but may operate independently of it. Despite the possibility that the perspective of the learner diverges from the evolved function of the learned actions, we do not propose that emic explanations are without purpose or should be dismissed. On the contrary, as pointed out across previous sections, the causal opacity of an action sequence as well as its resolvability are embedded in emic interpretations (they are products of the learner's mind) and hence it stands to reason that both explanatory levels need to be considered in order to account for the high fidelity transmission that is characteristic of traditional practices.

BST has the potential to generate many novel predictions about quasi-instrumental rituals by investigating how the presence and number of different cues modulate their frequency in a population. In some doctrinal religions, 'little traditions' based around quasi-instrumental rituals become widespread (Redfield, 1955). There is some evidence that although doctrinal traditions may cultivate strong forms of affiliation with the religious community, commensurate with their focus on the ritual stance, little traditions serve more closely the interests of individuals and kin groups seeking to ward off misfortune (Stanford and Whitehouse, 2021). Quasi-instrumental practices associated with little traditions may also spread readily across the boundaries of the doctrinal traditions in which they originate. BST would predict that magical practices crossing a group boundary are likely to undergo modification (because they are more likely to be copied by outsiders via an instrumental stance) while the transmission of the same magical practice within a group would entail higher fidelity copying (because it is part of the conventional repertoire that is maintained via affiliative motivations). This would account for the tendency for magical practices to contribute to group identity (e.g. the 'houses' associated with the mai or pai de santo in Brazilian candomblé traditions) while also allowing them to spread more freely across groups due to their goal salience (Whitehouse, 2011, 2021), which heightens the probability that they are transmitted via the instrumental stance.

3. The Psychological Foundations of BST

Much psychological research on social learning has been concerned mainly with the transmission of instrumental skills (Kenward, Karlsson, & Persson, 2011; Keupp, Behne, & Rakoczy, 2013; Lyons, Damrosch, Lin, Macris, & Keil, 2011; Nielsen, 2006b; Nielsen & Blank, 2011; Nielsen & Tomaselli, 2010; Over & Carpenter, 2009), thus potentially missing a wide range of behaviours that are relevant to our understanding of cultural evolution. We therefore turn to an outline of empirical research directly motivated by BST, incorporating both instrumental and ritual transmission and showing that heightened fidelity of opaque action primarily occurs in social reparatory scenarios, for example when an actor is facing the threat of ostracism (Watson-Jones, et al. 2014; Watson-Jones, Legare, & Whitehouse, 2016).

3.1 The Psychological Evidence Supporting BST

The two functions of social learning (to acquire instrumentally useful information and to affiliate with others) as well as their links to differential copying strategies have been discussed in developmental science for some time (Over & Carpenter, 2013; Uzgiris, 1981). Here we seek to extend awareness of bifocal functionality by putting these discussions of social learning into a context where documented differences in copying fidelity can be linked to the salience of social and action-based cues (see also Figure 1, above) that promote either the wish to affiliate via a norm-focused ritual stance or to achieve an end-state via a goal-focused instrumental stance (Herrmann et al., 2013; Legare et al., 2015; Watson-Jones et al., 2014). Although BST is grounded in anthropological theories of social learning (Whitehouse, 2011), it has been the object of a series of empirical investigations aimed at disambiguating the causes and effects of ritual and instrumental stances respectively using carefully controlled experiments (Herrmann et al., 2013; Legare et al., 2015; Watson-Jones et al., 2014; Watson-Jones, Whitehouse, & Legare, 2016). BST proposes that the differential functions of the two stances (ritual and instrumental; see Box 1.) favour a psychological division of labour where differing attentional and motivational states guide social learning in optimized ways. The ritual stance is prompted by irresolvable *causal opacity* and the absence of a salient goal or *goal demotion* (Box 1). In contrast, through the lens of an instrumental stance, actions are viewed as having a rational causal structure, each step contributing to an end goal. Accordingly, to be an efficient social learner, selectivity in the degree of copying fidelity is required. When learning a novel behaviour via action copying, we may rely largely on verbal or contextual cues that decide whether an action is best interpreted as conventional or instrumental (Clegg & Legare, 2016; Herrmann et al., 2013; Legare et al., 2015). Nonetheless, explicit cues in the form of verbal instructions are not always available (e.g. in the absence of pedagogy), so the action type itself often constitutes a source of either instrumentality or conventionality (Watson-Jones et al., 2014).If verbal or other extraneous contextual cues signal conventionality, then a learner should be sensitive to the action's social value and copy it with higher fidelity. Watson-Jones et al. demonstrate such a link by testing the psychological impact of witnessing third-party ostracism on copying fidelity in a sample of 3-6 year-olds when ritual versus instrumental

stances were made differentially salient. Children were shown video primes consisting of moving geometric figures reacting in an exclusionary (ostracism condition) or inclusionary manner (control condition) when approached by a new shape and were then encouraged to copy actions shown by a model. The sequences were either conventional, exhibiting the same start- and end-states (Figure 2), or instrumental, ending in a different end-state than the one they began with. Copying fidelity was highest for conventional actions in the *ostracism* prime condition. As predicted by BST, the attentional advantage that causally opaque action holds over its transparent counterparts in social settings is also exacerbated by directly experienced ostracism, as shown by follow-up designs (e.g. Watson-Jones et al., 2016) that manipulated group affiliation (ingroup/outgroup) using the *Cyberball* paradigm—a virtual ball tossing game in which participants are either included (base condition) or excluded (ostracism condition). Children who were subject to exclusion from their ingroup showed the highest copying fidelity of a causally opaque action, surpassing the copying accuracy of those who were previously included by their ingroup as well as those who were rejected or welcomed by an outgroup. It appears that the anxieties associated with social exclusion, even in a very mild form, have potentially far-reaching consequences for how modelled action is processed and attended to by social learners.

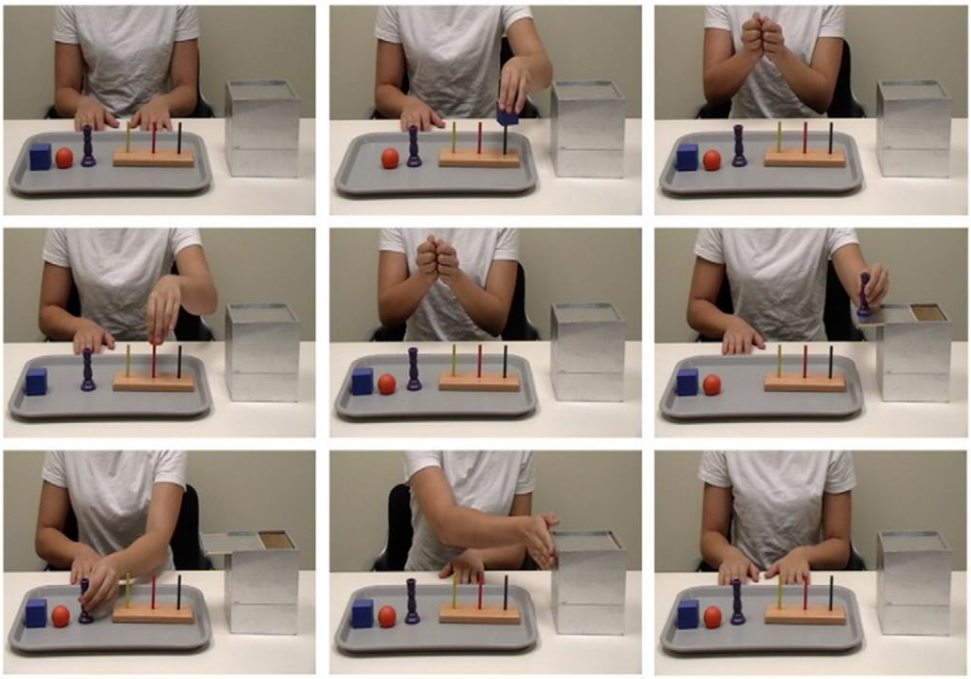

Figure 2. Start- and end-state equivalence, where item positions in the first step (first tile) of the sequence are identical to the end-state configuration (last tile), serves as conventional action condition (from Watson-Jones et al., 2014).

The studies reviewed above show that either directly or indirectly experienced negative interactions with others can cue changes in young learner's copying behaviour, demonstrating sensitivity to links between causally opaque action and favourable social outcomes. Such results are potentially consistent with the ritual stance being innate, domain-specific and rule guided (e.g. a biologically inherited module that is only active in social interactions and that follows a '*copy when ostracized*'-rule). Although such an account provides one possible framing (Whitehouse, 2021), we consider here the possibility that the observed data could just as plausibly be produced by domain-general processes such as associative learning, in which excitatory and inhibitory links between certain action types and social outcomes are forged during development (Heyes, 2012). For instance, the increased negative affect from social exclusion (rather than the exclusion itself) could result in additional efforts expended which, paired with the prediction errors that come from witnessed opacity, could result in overall higher fidelity scores for causally opaque action sequences but not causally transparent ones. Thus, an increase in non-specific arousal might have differential effects on ritual and instrumental copying, where heightened energetic investment in sequence copying could manifest in overall heightened fidelity, whereas putting more effort into endpoint copying does not result in as much accuracy gain given that the focus lies on the last step, the end goal. This potential alternative to a social reparatory account stresses the importance of efficient control conditions (e.g. general types of arousal as opposed to exclusion alone) in future BST designs to disentangle the potential specificity of stances from outputs that operate on mechanisms of general-purpose learning. If copying fidelity of opaque action is heightened in asocial conditions then domain-generality becomes a plausible interpretation.

Despite these outstanding questions that clearly show the importance of integrating BST into other fields of inquiry, the reviewed experiments are in line with much of the developmental literature and give rise to the reasonable assumption that at least a basic

understanding of action-reward links precedes the fully-fledged bifocal stance arrangement with which adults navigate society quite efficiently. Indeed, the literature documents a multitude of cases in which learning facilitates the acquisition and storage of social information such as norms, values and beliefs (Legare et al., 2012; Whitehouse, 2011). Norms are characterized as mutual agreements on how members of a group ought to conduct themselves in various social contexts and their acquisition marks a key milestone in development (Chudek &Henrich, 2011; Kenward, 2012; Kenward et al., 2011). Three-year olds intervene when witnessing norm violations through the use of normative language (Vaish, Missana, & Tomasello, 2011). Interestingly, children protest especially when an in-group member violates a norm but remain passive when witnessing transgressions committed by an unwitting outsider (Schmidt, Rakoczy, & Tomasello, 2012). Further, pretend play paradigms show how children are committed to arbitrary rules and the *proper* way to handle play scenarios (Nielsen, 2012; Rakoczyet al., 2009; Rakoczy, Warneken, & Tomasello, 2008). In Keupp et al. (2013), children protested more when a third party omitted irrelevant actions whenever those were labelled as conventional, pointing towards a normative interpretation of play behaviour. Similarly, Clegg and Legare (2016) demonstrate that linguistic cues that are indicative of convention rather than instrumentality boost copying fidelity of 3- to 6-year-olds in a necklace-making activity, further reinforcing the notion that the interpretation of the context in which knowledge is acquired, influences the way it is transmitted. Moreover, recent findings indicate that norm acquisition follows similar developmental trajectories across different societies (House et al., 2020). Overall, the reviewed findings paint a coherent picture of children's tendencies to attribute normativity when prompted, but instances of self-inferred norm attribution, occurring in contexts where it is not explicitly prompted, are also documented (Schmidt et al., 2016). Again, it is tempting to resort to cognitively rich explanations that accord children a fully-fledged understanding of norms, group affiliations and social dynamics; these findings are nonetheless suggestive that an early emerging rudimentary capacity to associate certain types of actions to certain types of outcomes lays the foundations for an appreciation of the affiliative value of cues signalling non-instrumentality. Accordingly, the ritual stance is geared towards obtaining social rewards, while the focus on reaping inanimate benefits within the context of goal achievement marks the instrumental stance (Herrmann et al., 2013; Legare et al., 2015; Watson-Jones et al., 2014).

In addition to studies tailored to test BST and normative protest, basic research on 'overimitation' (Box 1)(Call, Carpenter, & Tomasello, 2005; Horner & Whiten, 2005; Nagell, Olguin, & Tomasello, 1993; Whiten, Custance, Gomez, Teixidor, & Bard, 1996)—provides further support for our framework. While nonhuman primates sometimes copy simple instrumental behaviours, giving rise to localized tool-using cultures in some natural settings (Boesch & Boesch-Achermann, 2000), humans copy not only technically useful behaviours but also arbitrary social conventions to an extent unmatched by other primates (Horner & Whiten, 2005). Since at least the late Palaeolithic, we have produced and passed down increasingly complex cultural traditions, building cumulatively over generations (Hodder, 2012). This appears to have resulted from a subtle change in human learning strategies whereby instead of copying pragmatically any behaviour that produces desired outcomes, we also began to copy behaviour for which no causal rationale was available. Studies of the phenomenon have been carried out with western samples of both adults and children (Flynn & Smith, 2012; McGuigan, Gladstone, & Cook, 2012; McGuigan, Makinson, & Whiten, 2011; Whiten et al., 2016), as well as in a broad range of other cultural groups, from hunter-gatherers in Africa to swidden horticulturalists in Melanesia (Clegg & Legare, 2016a; Corriveau et al., 2017; Nielsen, Mushin, Tomaselli, & Whiten, 2014; Nielsen & Tomaselli, 2010; Taniguchi & Sanefuji, 2017).

Studies of overimitation often present research participants with tasks involving the extraction of a reward from a puzzle box following a demonstration of both causally transparent and causally opaque behaviours by the model (e.g. Horner & Whiten, 2005). In some cases, causal opacity is achieved by the insertion of behaviours that do not help with the extraction of the reward (e.g. "tapping the side of the jar with a feather and then unscrewing the lid"; Lyons et al., 2007). In other experiments, such behaviour does contribute to the end goal but only inefficiently (e.g. pushing a light switch with one's forehead rather than using one's hands; Buttelmann, Carpenter, Call, & Tomasello, 2007; Nagell et al., 1993). And in yet other designs, a more efficient method is available but overlooked by the model (Corriveau et al., 2017). In all such cases, human participants in the experiments copy the unnecessary or inefficient behaviour rather than prioritizing only the causally relevant aspects. Accordingly, BST proposes a much more pivotal role of non-

824 instrumental imitative strategies, by drawing attention to its ability to act as signals of
825 sharedness and aligned intention between actor and imitator (Over & Carpenter, 2012;
826 Tomasello, Carpenter, Call, Behne, & Moll, 2005). For instance, using a puzzle box
827 overimitation paradigm, Nielsen and Blank (2011) have shown that children tend to not
828 copy the causally irrelevant parts of a solution when the demonstrator is absent during
829 copying. Further, when previously primed by either a mimicry game or an instrumental
830 game, 2-year-olds overimitate or engage in selective copying respectively (Yu & Kushnir,
831 2014). Most importantly the effects of the mimicry prime only persist if the experimenter is
832 the same across both phases (thus excluding simple motor priming as an explanation).
833 Additionally 2-year-olds were also found to make use of unnecessary tools only in conditions
834 where the model is physically present as opposed to shown on a screen (Nielsen, Simcock, &
835 Jenkins, 2008). When learners copy observed behaviour accurately, they temporally
836 increase similarities between themselves and the model, which in turn increases their
837 likelihood of being perceived favourably by the model as shown by research documenting
838 that children preferentially help those that imitated them previously (Carpenter, Uebel, &
839 Tomasello, 2013; Harriet Over, Carpenter, Spears, & Gattis, 2013). Increased copying fidelity
840 observed in scenarios of overimitation has the effect of conserving shared conventions in
841 the cultural repertoire. This is in line with BST's core premise that the copying of causally
842 irrelevant elements is due to the motivation to affiliate with the model. There are various
843 potential points of proximate origin from which such motivations can spring. For instance,
844 social learners might retrospectively pick up a link between opaque copying and the
845 presence of social rewards, in a way that the behaviour becomes habitual and thus persists
846 even in the absence of further incentivisation. A learning individual may also extrapolate an
847 observed relationship between causally opaque elements and affiliative gains onto other
848 exemplars of ritual behaviour (in a prospective fashion). Most likely both retrospective and
849 prospective learning modes are involved in the ontogeny of BST, which is discussed in more
850 detail further below.

851 We now turn to accounts of overimitation that potentially compete with BST. These suggest
852 that non-instrumental copying is a mere by-product of our strong reliance on culture to
853 generate technological solutions. Overall, BST provides an alternative view to such by-
854 product explanations which attribute little to no importance to the type of action that is

855 transmitted. Moreover, we argue that BST provides a more encompassing framework within
856 which to understand the broad range of research findings on social learning.
857

858 ***3.2 By-product Accounts of Non-instrumental Copying***

859 One possible explanation for replication of instrumentally superfluous action might be the
860 learner's over-estimation of its causal efficacy (Lyons, Young, & Keil, 2007; Lyons et al.,
861 2011). When children copy causally opaque behaviour, it could be because they think the
862 model is trying to help them to acquire useful information and skills, even if it is not
863 immediately obvious how or why (Buchsbaum, Gopnik, Griffiths, & Shafto, 2011; Hoehl,
864 Zettersten, Schleihauf, Grätz, & Pauen, 2014; Vredenburgh, Kushnir, & Casasola, 2015). For
865 example, an influential early study on this topic revealed that infants were willing to copy
866 unnecessarily tortuous methods of activating a light switch (using the forehead instead of
867 the hands) after observing an adult model doing it that way (Gergely, Bekkering, & Király,
868 2002). The infants in this study were, however, sensitive to clues as to why the adult was
869 behaving so strangely. Thus, if the model was tightly wrapped in a blanket to constrain any
870 arm movements, the infants inferred that this was why she used her forehead in preference
871 to her hands to activate the switch and so they were more likely to use their hands to
872 activate the light rather than to copy the more inefficient method modelled to them. This
873 has led some psychologists to conclude that we only copy causally opaque behaviour on the
874 assumption that it must have a sound rationale, even if that is not apparent to us right
875 away. It has been suggested that such causal misattribution can be an implicit process, for
876 instance when children given training to identify correctly the causally unnecessary
877 elements in a modelled procedure still proceed to copy the irrelevant actions (Lyons et al.,
878 2011). In these studies, children are actively encouraged to recognise that the actions are
879 unnecessary and discouraged from reproducing them and yet they still engage in
880 overimitation. A possible explanation for this is 'automatic causal encoding' – that is, the
881 copying occurs outside conscious control (Lyons, Young, & Keil, 2007). In other words, when
882 children copy the unnecessary actions, they are motivated by implicit expectations of causal
883 structure and efficacy without being aware of it. These interpretations assume that non-
884 instrumental copying is merely a by-product of instrumental copying—*hyperactive* causal
885 attribution which automatically over-ascribes instrumental goals to all goal-demoted and

causally opaque actions. These accounts assume that the costs of falsely ascribing instrumental efficiency to functionally useless actions are offset by instances where those actions turn out to be teleologically efficacious.

Misattribution accounts are not without their critics (Kenward, 2012; Kenward et al., 2011).In line with BST, Kenward et al. (2011) propose that overimitation, rather than stemming from distorted causal beliefs, acts as facilitator in the development of norm acquisition. They found that when operating a puzzle box, children (age 4-5) copy unnecessary actions despite verbally declaring their inconsequentiality to reward extraction. Answers that signal either unquestioned copying or deontic reasoning were given upon further prompting, suggesting a normative mindset that utilizes opacity to achieve social rewards.

Considering these findings, it may be tempting to regard these two interpretations of overimitation – automatic causal encoding versus affiliative motivation – as mutually exclusive alternatives. That is, the copying of causally opaque behaviour is *either* motivated by expectations of hidden causal structure based on the imputation of pedagogic motivations (the 'instrumental stance') *or* by expectations of convention-learning based on imputation of normative motivations (the 'ritual stance'). Instead BST proposes that humans oscillate between stances during social learning, responding flexibly to the kinds of cues available to them when they observe modelled behaviour. Accordingly, "when the model is thought to have expertise in the performance of a task (signalled by cues like confidence, experience, success, and authority) one may anticipate an opportunity to learn something of practical use about the affordances of objects and so adopt an *instrumental stance* on the behaviour. But when the model is thought to be exemplifying a 'proper' or normative way of behaving (signalled by cues relevant to affiliation, conformism, or deference to tradition) it may be more appropriate to adopt a *ritual stance* on the behaviour and assume that it is simply the correct or 'done' way rather than the most causally efficacious way, of acting" (Whitehouse, 2021: 35).

By-product reasoning also features in the Action Parsing Model as applied to ritual (Lienard & Boyer, 2006; Liénard & Lawson, 2008). Action parsing encompasses three levels on which observed actions are processed: The *Script* level, which describes the overarching theme of

the action sequence (e.g. tidying the kitchen), *Behaviour*, which is a direct and end-focused description of the action (e.g. cleaning a glass), and *Gesture*, which describes the discrete sub-components of the behaviour (e.g. raising the glass, rubbing it with a cloth, etc.; Kapitány & Nielsen, 2017). Advocates of this approach argue that if instrumental outcomes of an action sequence are prevented by a missing causal link within the sequence (e.g. the glass and cloth are both raised but not brought into contact with each other during performance of a rubbing motion) then the attention of the observer shifts down to the gestural level. Accordingly, when witnessed, causal opacity and goal demotion cause a prediction error that comes from failing to parse the observed sequence on the level of *Behaviour*, thus promoting 'cognitive capture', which is characterized as increased allocation of attentional resources towards gestural components (Kapitány & Nielsen, 2017; Lienard & Boyer, 2006).This is in line with 'predictive coding' accounts, proposing that the discrepancies between top-down predictions and bottom-up sensory stimulation result in recruitment of additional neural resources geared towards updating the internal model of the agent and improving the accuracy of future estimates (Baldeweg, 2006; Friston & Kiebel, 2009; Huang & Rao, 2011). Predictions based on expected causality are therefore violated by goal absence, requiring ambiguity resolution. Indeed, the attention-grabbing nature of causal opacity improves recall (Kapitány, Kavanagh, Whitehouse, & Nielsen, 2018) and objects linked to opaque treatments are perceived as more special and desirable than occurring in causally transparent sequences (Kapitány & Nielsen, 2015). Causally opaque actions may make particularly unique (and hence efficient) identity markers precisely because they allow much greater combinatorial freedom on the gestural level rather than being confined to the realm of goal-directed *behaviour* anticipated by the observer.

The proposition that "steps become goals", that in causally opaque action sequences the absence of causal structure produces errors within the observer's mental prediction system that shifts attention away from the endpoint towards the sequence constituents, finds some limited support in developmental research. For instance, in one study (Carpenter, Call, &Tomasello, 2005), 12-month olds observed pretend play scenarios performed by adults (making a toy mouse hop across a mat with sound effects) either with a salient endpoint (the mouse was placed in a toy house) or not (it remained on the mat). Copying of play sequences was found to be more accurate (hopping and sounds were reproduced) in the

latter condition, suggesting that the absence of a goal prompted children to focus on the fine-grained level of gestural components. In similar designs, the level of fidelity with which young children copy decreases in conditions where they understand the purpose of the sequence (Williamson &Markman, 2006). Increased attention attributed to causally opaque action can therefore partially account for the heightened copying fidelity with which conventions and rituals are reportedly transmitted. Thus, the structure of an action sequence itself could drive observed differences in copying fidelity. As with explanations of casual misattribution given in the context of overimitation, the action parsing account would propose that transmission of causally irrelevant action is a by-product of error prediction systems that evolved for instrumental purposes. However, as argued by BST, which considers the tremendous social value of ritual participation(Whitehouse & Lanman, 2014; Xygalatas et al., 2013b), this is unlikely to be the end of the story. For instance, the Action Parsing Model fails to predict differences in copying fidelity primed by ostracism as foreseen by BST (and as reviewed above). Nothing about the level at which an action is parsed and attended to indicates why it should be more suitable for affiliation than some other type of action. Ultimately, we need BST to explain this type of modulation, where copying fidelity is not exclusively driven by attentional processes on the level of parsing (the cognitively "parasitic" nature of opacity and goal demotion) but also by the learner's social concerns and motivations during the process of transmission. More generally, we argue that the value of BST lies in its ability to add explanatory depth and nuance to the existing literature, while remaining consistent with the accounts described above: Causal misattributions, prediction errors, and the level of parsing, are all likely contributors to the high-fidelity transmission of conventions and traditions but none of them is sufficient to explain observed differences between ritual and instrumental learning.

In addition to its relevance to various sub-fields of social learning and cultural evolution research, another hallmark feature of BST is its emphasis on the importance of including different action types in experimental paradigms of social learning in order to disentangle its underlying motivations. In what follows, we make the case for using fine-grained distinctions between action characteristics in future research, stemming from a new framework of cultural action.

4 BST and the Cultural Action Framework

To identify types of action elements that cue a certain stance (instrumental or ritual)we propose a Cultural Action Framework (Figure 3), in which all observed actions can potentially be classified by their unique pathways and can thus be studied more efficiently by cognitive and social scientists alike. Bifurcations in our framework are based on the research reviewed above, suggesting that each step depends on the presence or absence of social cues that make one or other of these pathways more salient at any given moment. If an action is observed for which no end goal is discernible, we view it as irretrievably causally opaque thus activating the ritual stance and prompting high-fidelity copying (since any aspect of the action could be socially salient, e.g. as a stipulated requirement for functioning as a group identity marker, one must heed and copy everything). By contrast, if an observed action does have a discernible end-goal, this admits of two possible scenarios. In one scenario the means to this end-goal is causally transparent and resolvable, activating the instrumental stance and prompting less faithful transmission (since what mostly matters in this case is to achieve the outcome by whatever method works), leaving room for innovation to occur. An alternative scenario, however, is that despite the presence of an end-goal the sequence is causally opaque. If that opacity is perceived as resolvable, it activates the instrumental stance, again prompting lower-fidelity copying. If it is unresolved, but nevertheless seen as resolvable in principle, it would motivate the 'copy all, correct later' strategy that is typical of instrumental overimitation. The latter pathway is exemplified by "magical" practices which focus on an end-goal yet underspecify the causal pathway leading to the desired outcome (see above).

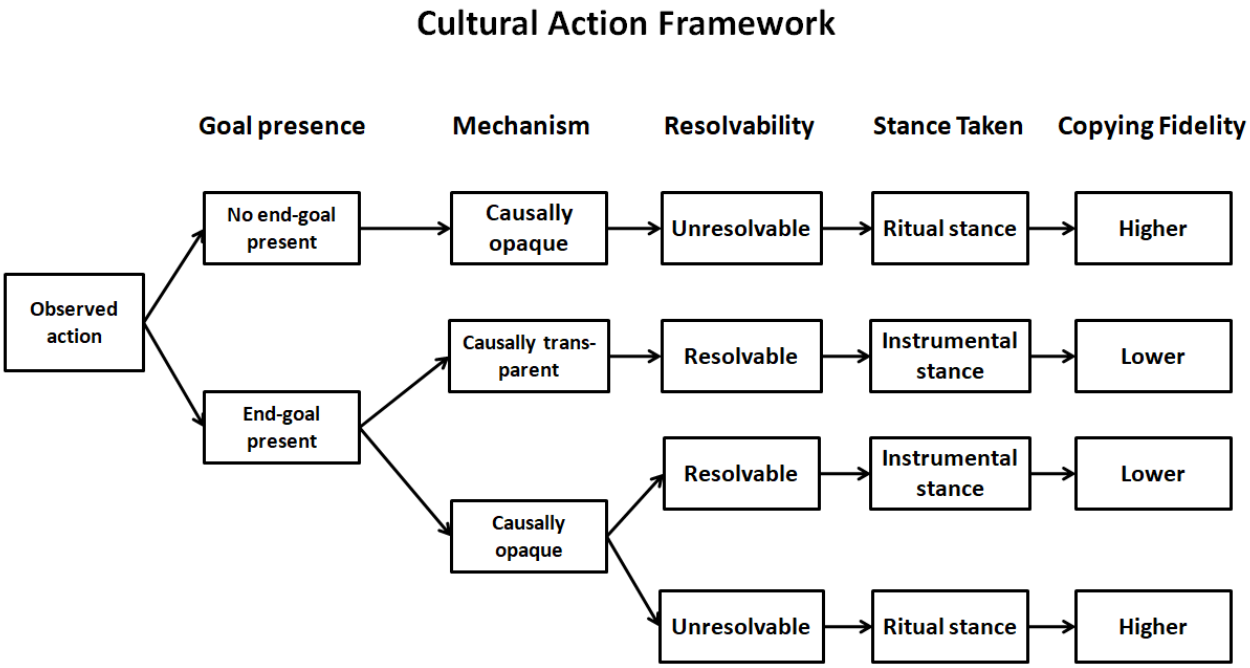

Figure 3. Cultural action framework based on goal-demotion, causal opacity and resolvability of opacity

This framework reveals the lop-sidedness of the current experimental literature. The vast majority of studies focus on one path: actions that are causally transparent and have a clear end-goal (e.g. Caldwell & Millen, 2010; Muthukrishna et al., 2013).Other studies that integrate causal opacity into their designs lack clear distinctions with regard to varying degrees of resolvability. The proposed framework therefore aims to disambiguate the cognitive processes giving rise to a multitude of cultural traits and to facilitate more fine-grained analyses of cultural transmission. For instance, *microsociety* designs make use of experimental diffusion chains where an action is copied from one person to the next to track how it changes over multiple intergenerational transmissions (Caldwell et al., 2018). BST predicts that under conditions where social cues are salient (e.g. presence of ostracism threat), the modelling of goal-demoted and irretrievably opaque content prompts greater resistance to deviation than instrumental and resolvable action. More specifically, this could be shown by modifying already popular experiments (e.g. using paper planes; Caldwell & Millen, 2008) by including causally opaque elements (such as particular types of folds that do not contribute to the plane's aerodynamics or superfluous gestures during construction) and then track the longevity of these elements compared to their functional equivalents.

Thus, by capturing cultural elements that are beyond the realm of technology, existing research paradigms can be refined by the current framework in ways that allow for broader conclusions about cultural evolution. As we will now discuss, investigations that are attentive to the differential transmission patterns of various cultural action types constitute fertile soil for future research.

5 The Future of BST

By providing explanations for the co-existence of copying fidelity and innovative change, as well as an alternative framework to by-product accounts of overimitation, BST opens many new avenues for research on the psychological mechanisms and evolutionary-developmental origins of cultural learning. A key question about mechanisms asks whether the stances impact cultural learning by modulating attention, motivation, or – most likely – both. As the Action Parsing Model indicates (Lienard & Boyer, 2006; Liénard & Lawson, 2008), early processing of observed action and its context could trigger attention to molar or molecular features of ensuing behaviour, promoting more effective learning about coarse-grain (instrumental stance) or finer-grain (ritual stance) features of the observed action (Leighton, Bird, & Heyes, 2010). While the attainment of a desired terminal configuration during instrumental copying emphasizes coarse and molar features, the sequence endpoint of opaque action is less salient due to its causal inconsequentiality, thus shifting attention towards the molecular features of the action (i.e. “steps become goals”). In addition, features of the action and context could encourage the learner to anticipate social or nonsocial rewards for successful performance (e.g. approval and/or a sticker) and to work harder, producing higher fidelity copying, when social rewards are both anticipated and needed– for example, due to recent priming of ostracism threat(Watson-Jones et al., 2014, 2016).

To investigate these possibilities empirically, the contribution of attentional processes could be assessed with eye-tracking, and with variants of the overimitation and start-end-state paradigm using recognition tests. In these variants, children could be asked to identify the action they have just observed from a set of images that differ from the target action at the molar or molecular level. If the stances have their effects via attention as well as motivation,

more identification errors would occur when actions with a salient goal have been changed at the molecular level, and when actions without a salient goal have been changed at the molar level. In other words, if an observed behaviour is goal-oriented and thus encoded at the coarse and molar level ("tidying the kitchen"), it should be more difficult to detect changes at the level of its sequence constituents (detecting whether a glass was first "raised" and then "scrubbed" or vice versa). Conversely, the attentional shift towards the molecular level as is the case for ritualistic behaviour ("the candle is the first item to be placed on the table"), should render identification at the molar level more difficult.

The contribution of motivational processes, already established using the start-end-state procedure (Watson-Jones et al., 2016), could be probed further by reversal studies where, in a novel context, low-fidelity copying produces social rewards, and high-fidelity copying produces nonsocial rewards. Once they are familiar with this looking-glass world, do adults and children begin to copy with higher fidelity when they want nonsocial outcomes? Such a reversal of the typical pattern would provide strong evidence that copying fidelity is motivated by expectations about the type of reward that is likely to follow. Similarly, looking-glass paradigms can be used to investigate action-reward reversals, where for instance adopting the instrumental stance reaps social gains while irresolvable causal opacity is paired with inanimate rewards. Such designs can potentially uncover the motivational potency with which the anticipation of either reward modulates the fidelity of transmitted action.

Another important mechanism question asks whether the ritual and instrumental stances differentially recruit imitation and emulation. In imitation, narrowly defined, the observer copies body movements - the way that parts of the body move relative to one another (e.g. fist to chin) – whereas in emulation, the observer reproduces object movements (e.g. purple cube to red peg) (Heyes, in press; 2021; Heyes, 1993; M. Tomasello, Kruger, & Ratner, 1993). Given that many group-defining communicative and ritual actions are intransitive, consisting of gestures and postures that do not involve objects (such as rhythmic dancing, marching and more generally rituals that rely on joint and synchronous movements; Hove & Risen, 2009; Wiltermuth & Heath, 2009), it is likely that the ritual stance primes imitation more strongly than the instrumental stance. This would be significant because there is a substantial body of evidence that, unlike emulation, imitation

involves distinctively human cognitive processes that are specialised for cultural learning (Tennie et al., 2009; Whiten, 2017). Previous experiments have not indicated a bias towards imitation in the ritual stance, possibly because the object movements in both conventional and instrumental sequences were highly salient compared with the body movements (Watson-Jones et al. 2014; 2016). Experiments designed to probe the roles of imitation and emulation would use action sequences in which body movements and object movements are equal in frequency, distribution and salience, predicting higher fidelity copying of the body movements in conventional than in instrumental conditions.

More broadly, BST encourages research of a kind that is often neglected in developmental and evolutionary psychology; investigating the extent to which a focal behavioural competence depends on, rather than scaffolds, cognitive development. At the dependence end of this spectrum, stance behaviour–higher fidelity conventional rather than instrumental copying–may be deliberative (see Section 2.3.3). Before copying, the actor categorises the action they observed as conventional or instrumental; deliberates about the model's personal and communicative intentions; and reflects on their own desires to achieve affiliative rather than instrumental goals. In other words, stance behaviour may require a learner, even a young child, to have a range of concepts – including 'convention' or 'norm', 'cause', 'should' and 'intention' – and to be capable of making inferences using these concepts in ways that are typically described as normative reasoning, causal reasoning, and mentalising. In contrast, at the scaffolding end of the spectrum, a child may begin to show stance behaviour based on simple knowledge of contingencies. She may have noticed that high fidelity copying is more likely to meet with social approval when the action and context have certain observable features, ABC, than when they have other features, XYZ. Consequently, without knowing why these contingencies hold – without understanding convention or instrumentality – she expects high fidelity copying to produce richer social rewards under conditions ABC than conditions XYZ and acts accordingly (Grusec &Abramovitch, 1982; Young, Krantz, McClannahan, &Poulson, 1994). Thus, at the scaffolding end of the continuum, stance behaviour has humble beginnings, but it provides a crucial platform for the development of complex normative and causal cognition. Once the child can distinguish ABC from XYZ – once these two sets of observable cues are distinct in her mind – she can begin to build the concepts through which the two categories are

understood by adult members of her culture. For example, she can connect adults' statements about what 'we do' to ABC, and statements about what 'works' to XYZ, and through these build concepts of convention and causation, and capacity for normative reasoning.

Another, related question concerns the way in which the stances are inherited, genetically or culturally (see Section 2.2). On the one hand, bifocal stance psychology could be an 'innate module' or 'cognitive instinct' (Pinker, 2010; Tooby & Cosmides, 2005). In other words, it is possible that bifocal stance psychology evolved via natural selection operating on genetic variants, and that all (or nearly all) contemporary humans genetically inherit a specific propensity to develop the two stances; a propensity that is realised with minimal input from experience (e.g. only the opportunity to observe the behaviour of others). At the other end of the inheritance spectrum, bifocal stance psychology could be a 'cognitive gadget'(Box. 1); a product of natural selection operating on socially learned variants (Birch & Heyes, in press;Heyes, 2018). In the gadget case, the genetically inherited ingredients would be nonspecific (domain-general as opposed to domain-specific; Box 1).These ingredients may include only an enhanced capacity for detecting action-reward contingencies, and heightened social motivation compared with our primate ancestors, but the experiential requirements for the development of bifocal stance psychology would be much more specific. For example, it may be necessary for children to grow up among adults who consistently approve high fidelity copying of conventional actions, and who encourage active curiosity about how things work.

In principle, the two stances could be genetically or culturally inherited whether stance behaviour depends on, or scaffolds the development of, concepts and complex cognitive processes. However, given that stance behaviour emerges early in development (by the age of 3; Watson-Jones et al., 2014), when children have had limited opportunity to learn about normativity and other minds through conversation, high dependence on complex cognition would imply that bifocal stance psychology is at the genetic end, and high scaffolding would imply that it lies at the cultural end, of the inheritance spectrum. Consequently, the dependence and inheritance questions can be investigated together by examining co-variation between cultural learning experiences and the development of bifocal stance psychology – across individuals, groups, and cultures.

For example, do children who have been more consistently rewarded for high fidelity copying of conventional actions, and encouraged to 'have a go' with novel objects, develop stance behaviour sooner than children with less of this cultural learning experience? Do they develop a richer understanding of normativity indexed by verbal measures? Similarly, do children from more traditional cultures show signs of bifocal stance psychology sooner than children from less traditional cultures, and use different sets of cues to delineate conventional and instrumental behaviour? Affirmative answers to questions like these would suggest a powerful role for cultural learning in the development of bifocal stance psychology; that this cognitive specialisation for the cultural evolution of practical skills is itself a product of cultural evolution.

6 Concluding Remarks

No other species is as dependent on culture as humans. Cultural adaptations resulting from collectively accumulated bodies of knowledge turned a tropical primate into the ecologically dominant species on the planet. To explain why, we have argued for the Bifocal Stance Theory (BST) of cultural evolution, which proposes that the co-existence of innovation and adherence to tradition results from our ability to adopt different motivational stances and associated copying paradigms attuned to the different functional affordances of the behaviour to be transmitted. BST stands in stark contrast to popular by-product accounts, which assume that faithful copying of causally irrelevant action is due to reliance on instrumental copying alone. We have proposed empirically tractable alternative accounts of the evolutionary origins of the two stances and a novel cultural action framework to help ensure that experimental designs approximate more closely real-life cultural transmission. BST not only illuminates historical patterns of differential cultural diffusion (continuity and change in both technology and ritual life) but also raises new questions about the cognitive underpinnings of cultural evolution. Above all, we propose that if our reliance on culture is what makes humans stand out on the tree of life then studying the bifocal stances that enable culture may hold the key to understanding the evolutionary origins of human uniqueness.

Funding Statement: HW's work on this article was supported by an Advanced Grant from the European Research Council under the European Union's Horizon 2020 Research and Innovation Programme (Grant Agreement 694986).

Conflicts of Interest: None.

References

Apperly, I. (2010). *Mindreaders: the cognitive basis of" theory of mind"*. Psychology Press.

Aronson, E., & Mills, J. (1959). The effect of severity of initiation on liking for a group. *Journal of Abnormal and Social Psychology*. https://doi.org/10.1037/h0047195

Atkisson, C., O'Brien, M. J., & Mesoudi, A. (2012). Adult Learners in a Novel Environment Use Prestige-Biased Social Learning. *Evolutionary Psychology*, *10*(3), 147470491201000. https://doi.org/10.1177/147470491201000309

Baldeweg, T. (2006). Repetition effects to sounds: Evidence for predictive coding in the auditory system. *Trends in Cognitive Sciences*, *10*(3), 93–94. https://doi.org/10.1016/j.tics.2006.01.010

Birch, J., & Heyes, C. M. (n.d.). The cultural evolution of cultural evolution. *Philosophical Transactions of the Royal Society : B*.

Boesch, C., & Boesch-Achermann, H. (2000). *The chimpanzees of the Taï Forest: Behavioural ecology and evolution*. Oxford University Press.

Buchsbaum, D., Gopnik, A., Griffiths, T. L., & Shafto, P. (2011). Children's imitation of causal action sequences is influenced by statistical and pedagogical evidence. *Cognition*, *120*(3), 331–340. https://doi.org/10.1016/j.cognition.2010.12.001

Butler, J., Floccia, C., Goslin, J., Panneton, R., & Tech, V. (2011). *Infants ' Discrimination of Familiar and Unfamiliar Accents in Speech*. *16*(4), 392–417. https://doi.org/10.1111/j.1532-7078.2010.00050.x

Buttelmann, D., Carpenter, M., Call, J., & Tomasello, M. (2007). Enculturated chimpanzees imitate rationally. *Developmental Science*, *10*(4), F31–F38. https://doi.org/10.1111/j.1467-7687.2007.00630.x

Caldwell, C. A., & Millen, A. E. (2008). Experimental models for testing hypotheses about cumulative cultural evolution. *Evolution and Human Behavior*, *29*(3), 165–171. https://doi.org/10.1016/j.evolhumbehav.2007.12.001

Caldwell, C. A., & Millen, A. E. (2010). Human cumulative culture in the laboratory: Effects of (micro) population size. *Learning & Behavior*, *38*(3), 310–318. https://doi.org/10.3758/LB.38.3.310

Caldwell, C. A., Renner, E., & Atkinson, M. (2018). Human Teaching and Cumulative Cultural Evolution. *Review of Philosophy and Psychology*, *9*(4), 751–770. https://doi.org/10.1007/s13164-017-0346-3

Call, J., Carpenter, M., & Tomasello, M. (2005). Copying results and copying actions in the process of social learning: chimpanzees (Pan troglodytes) and human children (Homo sapiens). *Animal Cognition*, *8*(3), 151–163. https://doi.org/10.1007/s10071-004-0237-8

Carpenter, M., Call, J., & Tomasello, M. (2005). Twelve- and 18-month-olds copy actions in terms of goals. *Developmental Science*, *8*(1), F13–F20. https://doi.org/10.1111/j.1467-7687.2004.00385.x

Carpenter, M., Uebel, J., & Tomasello, M. (2013). Being mimicked increases prosocial behavior in 18-month-old infants. *Child Development*. https://doi.org/10.1111/cdev.12083

Catmur, C., & Heyes, C. (2013). Is it what you do, or when you do it? The roles of contingency and similarity in pro-social effects of imitation. *Cognitive Science*, *37*(8), 1541–1552.

Cavalli-Sforza, L. L., & Feldman, M. W. (1981). *Cultural Transmission and Evolution: A Quantitative Approach*. Chicago: Princeton University Press.

Chudek, M., & Henrich, J. (2011). Culture-gene coevolution, norm-psychology and the emergence of human prosociality. *Trends in Cognitive Sciences*. https://doi.org/10.1016/j.tics.2011.03.003

Clegg, J. M., & Legare, C. H. (2016a). A cross-cultural comparison of children's imitative flexibility. *Developmental Psychology*. https://doi.org/10.1037/dev0000131

Clegg, J. M., & Legare, C. H. (2016b). Instrumental and Conventional Interpretations of Behavior Are Associated With Distinct Outcomes in Early Childhood. *Child Development*, *87*(2), 527–542. https://doi.org/10.1111/cdev.12472

Corriveau, K. H., DiYanni, C. J., Clegg, J. M., Min, G., Chin, J., & Nasrini, J. (2017). Cultural differences in the imitation and transmission of inefficient actions. *Journal of Experimental Child Psychology*, *161*, 1–18. https://doi.org/10.1016/j.jecp.2017.03.002

Cosmides, L., & Tooby, J. (1994). Origins of domain specificity: The evolution of functional organization. In L. A. Hirschfeld & S. A. Gelman (Eds.), *Mapping the Mind: Domain Specificity in Cognition and Culture* (pp. 85–116). https://doi.org/10.1017/CBO9780511752902.005

Diesendruck, G., & Markson, L. (2001). Children's avoidance of lexical overlap : A pragmatic account. *Developmental Psychology*, *37*(May 2016), 630–641. https://doi.org/10.1037//0012-1649.37.5.630

Durkheim, E. (1965). *The Elementary forms of the religious life*. London: Free Press.

Evans, J. S. B. T., & Stanovich, K. E. (2013). Dual-Process Theories of Higher Cognition. *Perspectives on Psychological Science*, *8*(3), 223–241.

https://doi.org/10.1177/1745691612460685

Fessler, D. M., & Holbrook, C. (2016). Synchronized behavior increases assessments of the formidability and cohesion of coalitions. *Evolution and Human Behavior*, *37*(6), 502–509.

Flynn, E., & Smith, K. (2012). Investigating the Mechanisms of Cultural Acquisition. *Social Psychology*, *43*(4), 185–195. https://doi.org/10.1027/1864-9335/a000119

Friston, K., & Kiebel, S. (2009). Predictive coding under the free-energy principle. *Philosophical Transactions of the Royal Society B: Biological Sciences*, *364*(1521), 1211–1221. https://doi.org/10.1098/rstb.2008.0300

Gergely, G., Bekkering, H., & Király, I. (2002). Rational imitation in preverbal infants. *Nature*, *415*(6873), 755–755. https://doi.org/10.1038/415755a

Gergely, G., & Csibra, G. (2006). Sylvia's Recipe: The Role of Imitation and Pedagogy in the Transmission of Human Culture. In *Roots of Human Sociality: Culture, Cognition, and Human Interaction* (pp. 229–255).

Goupil, L., & Kouider, S. (2019). Developing a reflective mind: from core metacognition to explicit self-reflection. *Current Directions in Psychological Science*, *28*(4), 403–408.

Graham, S. A., Stock, H., & Henderson, A. M. (2006). Nineteen-Month-Olds ' Understanding of the Conventionality of Object Labels Versus Desires Nineteen-Month-Olds ' Understanding of the Conventionality of Object Labels Versus Desires. *Infancy*, *9*(3), 341–350.

Grusec, J. E., & Abramovitch, R. (1982). Imitation of Peers and Adults in a Natural Setting: A Functional Analysis. *Child Development*, *53*(3), 636. https://doi.org/10.2307/1129374

Hagen, E. H., & Bryant, G. A. (2003). Music and dance as a coalition signaling system. *Human Nature*, *14*(1), 21–51.

Hart, C. W. M., Pilling, A., & Goodale, J. (1988). *The Tiwi of North Australia (Case Studies in Cultural Anthropology)*. Wadsworth Publishing Co Inc.

Henrich, J. (2009). The evolution of costly displays, cooperation and religion. *Evolution and Human Behavior*, *30*(4), 244–260. https://doi.org/10.1016/j.evolhumbehav.2009.03.005

Herrmann, P. A., Legare, C. H., Harris, P. L., & Whitehouse, H. (2013). Stick to the script: The effect of witnessing multiple actors on children's imitation. *Cognition*, *129*(3), 536–543. https://doi.org/10.1016/j.cognition.2013.08.010

Heyes, C. (2018). *Cognitive gadgets: The cultural evolution of thinking*. Harvard University Press.

Heyes, C., Bang, D., Shea, N., Frith, C. D., & Fleming, S. M. (2020). Knowing ourselves together: The cultural origins of metacognition. *Trends in Cognitive Sciences*, *24*(5), 349–362.

Heyes, C. M. (n.d.). Imitation and culture: what gives? *Mind and Language*.

Heyes, C. M. (2021). Primer on "imitation." *Current Biology*, *31*, R215–R240.

Heyes, C.M. (1993). Imitation, culture and cognition. *Animal Behaviour*, *46*(5), 999–1010. https://doi.org/10.1006/anbe.1993.1281

Heyes, C.M. (2019). Précis of Cognitive Gadgets: The Cultural Evolution of Thinking. *Behavioral and Brain Sciences*, *42*, e169:1-5.

Heyes, Cecilia. (2012). What's social about social learning? *Journal of Comparative Psychology*, *126*(2), 193–202. https://doi.org/10.1037/a0025180

Heyes, Cecilia. (2018). Enquire within: cultural evolution and cognitive science. *Philosophical Transactions of the Royal Society B: Biological Sciences*, *373*(1743), 20170051. https://doi.org/10.1098/rstb.2017.0051

Heyes, Cecilia M., & Frith, C. D. (2014). The cultural evolution of mind reading. *Science*, *344*(6190). https://doi.org/10.1126/science.1243091

Hodder, I. (2012). *Entangled: An archaeology of the relationships between humans and things*. New Jersey: Wiley-Blackwell.

Hoehl, S., Zettersten, M., Schleihauf, H., Grätz, S., & Pauen, S. (2014). The role of social interaction and pedagogical cues for eliciting and reducing overimitation in preschoolers. *Journal of Experimental Child Psychology*. https://doi.org/10.1016/j.jecp.2013.12.012

Horner, V., & Whiten, A. (2005). Causal knowledge and imitation/emulation switching in chimpanzees (Pan troglodytes) and children (Homo sapiens). *Animal Cognition*, *8*(3), 164–181. https://doi.org/10.1007/s10071-004-0239-6

House, B. R., Kanngiesser, P., Barrett, H. C., Broesch, T., Cebioglu, S., Crittenden, A. N., … Yilmaz, S. (2020). Universal norm psychology leads to societal diversity in prosocial behaviour and development. *Nature Human Behaviour*, *4*(1), 36–44.

Hove, M. J., & Risen, J. L. (2009). It's All in the Timing: Interpersonal Synchrony Increases Affiliation. *Social Cognition*, *27*(6), 949–960. https://doi.org/10.1521/soco.2009.27.6.949

Huang, Y., & Rao, R. P. N. (2011). Predictive coding. *Wiley Interdisciplinary Reviews: Cognitive Science*, *2*(5), 580–593. https://doi.org/10.1002/wcs.142

Humphrey, C., & Laidlaw, J. (1994). *The Archetypal Actions of Ritual: A Theory of Ritual Illustrated by the Jain Rite of Worship*. Clarendon Press.

Irons, W. (2001). Religion as a Hard-to-Fake-Sign of Commitment. *Evolution and the Capacity for Commitment*, 292–302.

Kahneman, D. (2003). A perspective on judgment and choice: mapping bounded rationality. *Am. Psychol.*, (58), 697–720.

Kapitány, R., Kavanagh, C., Whitehouse, H., & Nielsen, M. (2018). Examining memory for ritualized gesture in complex causal sequences. *Cognition*, *181*, 46–57. https://doi.org/10.1016/j.cognition.2018.08.005

Kapitány, Rohan, & Nielsen, M. (2015). Adopting the ritual stance: The role of opacity and context in ritual and everyday actions. *Cognition*, *145*, 13–29. https://doi.org/10.1016/j.cognition.2015.08.002

Kapitány, Rohan, & Nielsen, M. (2017). The ritual stance and the precaution system: the role of goal-demotion and opacity in ritual and everyday actions. *Religion, Brain & Behavior*, *7*(1), 27–42. https://doi.org/10.1080/2153599X.2016.1141792

Kapitány, Rohan, & Nielsen, M. (2019). Ritualized Objects: How We Perceive and Respond to Causally Opaque and Goal Demoted Action. *Journal of Cognition and Culture*, *19*(1–2), 170–194. https://doi.org/10.1163/15685373-12340053

Kenward, B. (2012). Over-imitating preschoolers believe unnecessary actions are normative and enforce their performance by a third party. *Journal of Experimental Child Psychology*, *112*(2), 195–207. https://doi.org/10.1016/j.jecp.2012.02.006

Kenward, B., Karlsson, M., & Persson, J. (2011). Over-imitation is better explained by norm learning than by distorted causal learning. *Proceedings of the Royal Society B: Biological Sciences*, *278*(1709), 1239–1246. https://doi.org/10.1098/rspb.2010.1399

Kertzer, D. (1988). *Ritual, Politics, and Power*. New Haven, CT: Yale University Press.

Kertzer, D. I. (1989). *Ritual, politics, and power*. New Haven, CT: Yale University Press.

Keupp, S., Behne, T., & Rakoczy, H. (2013). Why do children overimitate? Normativity is crucial. *Journal of Experimental Child Psychology*, *116*(2), 392–406. https://doi.org/10.1016/j.jecp.2013.07.002

Khaldūn, I. (1958). *The Muqaddimah : An introduction to history. Translated from the Arabic by Franz Rosenthal. 3 vols.* New York: Princeton.

Kinzler, K. D., Corriveau, K. H., & Harris, P. L. (2011). Children's selective trust in native-accented speakers. *Developmental Science*, *14*(1), 106–111. https://doi.org/10.1111/j.1467-7687.2010.00965.x

Kinzler, K. D., Shutts, K., Dejesus, J., & Spelke, E. S. (2009). Accent trumps race in guiding children's social preferences. *Social Cognition*. https://doi.org/10.1521/soco.2009.27.4.623

Koenig, M. A., Cle, F., & Harris, P. L. (2004). Trust in Testimony Children ' s Use of True and False Statements. *Psychological Science*, *15*(10), 694–698.

Lakin, J. L., & Chartrand, T. L. (2003). Using nonconscious behavioral mimicry to create affiliation and rapport. *Psychological Science*, *14*(4), 334–339.

Lawrence, P. (1989). *Road Belong Cargo: A Study of the Cargo Movement in the Southern Madang District, New Guinea*. Waveland Pr Inc.

1357 Legare, C. H., Evans, E. M., Rosengren, K. S., & Harris, P. L. (2012). The Coexistence of
1358 Natural and Supernatural Explanations Across Cultures and Development. *Child*
1359 *Development*, *83*(3), 779–793. https://doi.org/10.1111/j.1467-8624.2012.01743.x

1360 Legare, C. H., & Herrmann, P. A. (2013). Cognitive consequences and constraints on
1361 reasoning about ritual. *Religion, Brain & Behavior*, *3*(1), 63–65.
1362 https://doi.org/10.1080/2153599X.2012.736710

1363 Legare, C. H., & Souza, A. L. (2014). Searching for control: Priming randomness increases the
1364 evaluation of ritual efficacy. *Cognitive Science*, *38*(1), 152–161.
1365 https://doi.org/10.1111/cogs.12077

1366 Legare, C. H., Wen, N. J., Herrmann, P. A., & Whitehouse, H. (2015). Imitative flexibility and
1367 the development of cultural learning. *Cognition*, *142*(512), 351–361.
1368 https://doi.org/10.1016/j.cognition.2015.05.020

1369 Legare, C., & Souza, A. L. (2012). Evidence from the supernatural. *Cognition*, *124*(1), 1–15.

1370 Leighton, J., Bird, G., & Heyes, C. (2010). 'Goals' are not an integral component of imitation.
1371 *Cognition*, *114*(3), 423–435. https://doi.org/10.1016/j.cognition.2009.11.001

1372 Lewis, I. M. (2002). *Ecstatic religion: a study of shamanism and spirit possession*. Routledge.

1373 Lienard, P., & Boyer, P. (2006). Whence Collective Rituals? A Cultural Selection Model of
1374 Ritualized Behavior. *American Anthropologist*, *108*(4), 814–827.
1375 https://doi.org/10.1525/aa.2006.108.4.814

1376 Liénard, P., & Lawson, E. T. (2008). Evoked culture, ritualization and religious rituals.
1377 *Religion*, *38*(2), 157–171. https://doi.org/10.1016/j.religion.2008.01.004

1378 Lu, C. H., & Proctor, R. W. (1995). The influence of irrelevant location information on
1379 performance: A review of the Simon and spatial Stroop effects. *Psychonomic Bulletin &*
1380 *Review*, *2*(2), 174–207.

1381 Lyons, D. E., Young, A. G., & Keil, F. C. (2007). The hidden structure of overimitation.
1382 *Proceedings of the National Academy of Sciences*, *104*(50), 19751–19756.
1383 https://doi.org/10.1073/pnas.0704452104

1384 Lyons, Derek E., Damrosch, D. H., Lin, J. K., Macris, D. M., & Keil, F. C. (2011). The scope and
1385 limits of overimitation in the transmission of artefact culture. *Philosophical*
1386 *Transactions of the Royal Society B: Biological Sciences*, *366*(1567), 1158–1167.
1387 https://doi.org/10.1098/rstb.2010.0335

1388 MacLeod, C. M. (1991). Half a century of research on the Stroop effect: an integrative
1389 review. *Psychological Bulletin*, *109*(2), 163.

1390 McGuigan, N., Gladstone, D., & Cook, L. (2012). Is the Cultural Transmission of Irrelevant
1391 Tool Actions in Adult Humans (Homo Sapiens) Best Explained as the Result of an
1392 Evolved Conformist Bias? *PLoS ONE*, *7*(12), e50863.
1393 https://doi.org/10.1371/journal.pone.0050863

1394 McGuigan, N., Makinson, J., & Whiten, A. (2011). From over-imitation to super-copying:

Adults imitate causally irrelevant aspects of tool use with higher fidelity than young children. *British Journal of Psychology*. https://doi.org/10.1348/000712610X493115

McKay, R., & Whitehouse, H. (2015). Religion and morality. *Psychological Bulletin*, *141*(2), 447–473. https://doi.org/10.1037/a0038455

Mehr, S. A., Krasnow, M. M., Bryant, G. A., & Hagen, E. H. (2020). Origins of music in credible signaling. *Behavioural and Brain Sciences*, 1–41.

Mesoudi, A. (2008). An experimental simulation of the "copy-successful-individuals" cultural learning strategy: adaptive landscapes, producer–scrounger dynamics, and informational access costs. *Evolution and Human Behavior*, *29*(5), 350–363. https://doi.org/10.1016/j.evolhumbehav.2008.04.005

Mesoudi, A. (2011). An experimental comparison of human social learning strategies: payoff-biased social learning is adaptive but underused. *Evolution and Human Behavior*, *32*(5), 334–342. https://doi.org/10.1016/j.evolhumbehav.2010.12.001

Mesoudi, A., & Thornton, A. (2018). What is cumulative cultural evolution? *Proceedings of the Royal Society B: Biological Sciences*, *285*(1880), 20180712. https://doi.org/10.1098/rspb.2018.0712

Mulak, K. E., Best, C. T., Tyler, M. D., Kitamura, C., & Irwin, J. R. (2014). Development of phonological constancy: 19-month-olds, but not 15-month-olds, identify words in a non-native regional accent. *Child Development*, *84*(6), 2064–2078. https://doi.org/10.1111/cdev.12087.Development

Muthukrishna, M., Shulman, B. W., Vasilescu, V., & Henrich, J. (2013). Sociality influences cultural complexity. *Proceedings of the Royal Society B: Biological Sciences*, *281*(1774), 20132511. https://doi.org/10.1098/rspb.2013.2511

Nagell, K., Olguin, R. S., & Tomasello, M. (1993). Processes of social learning in the tool use of chimpanzees (Pan troglodytes) and human children (Homo sapiens). *Journal of Comparative Psychology*, *107*(2), 174–186. https://doi.org/10.1037/0735-7036.107.2.174

Nielsen, M. (2006). Copying actions and copying outcomes: Social learning through the second year. *Developmental Psychology*. https://doi.org/10.1037/0012-1649.42.3.555

Nielsen, M. (2012). Imitation, pretend play, and childhood: Essential elements in the evolution of human culture? *Journal of Comparative Psychology*, *126*(2), 170–181. https://doi.org/10.1037/a0025168

Nielsen, M., & Blank, C. (2011). Imitation in young children: When who gets copied is more important than what gets copied. *Developmental Psychology*, *47*(4), 1050–1053. https://doi.org/10.1037/a0023866

Nielsen, M., Mushin, I., Tomaselli, K., & Whiten, A. (2014). Where Culture Takes Hold: "Overimitation" and Its Flexible Deployment in Western, Aboriginal, and Bushmen Children. *Child Development*. https://doi.org/10.1111/cdev.12265

Nielsen, M., Simcock, G., & Jenkins, L. (2008). The effect of social engagement on 24-month-olds' imitation from live and televised models. *Developmental Science*. https://doi.org/10.1111/j.1467-7687.2008.00722.x

Nielsen, M., & Tomaselli, K. (2010). Overimitation in kalahari bushman children and the origins of human cultural cognition. *Psychological Science*. https://doi.org/10.1177/0956797610368808

Nielsen, M., Tomaselli, K., & Kapitány, R. (2018). The influence of goal demotion on children's reproduction of ritual behavior. *Evolution and Human Behavior*, *39*(3), 343–348. https://doi.org/10.1016/j.evolhumbehav.2018.02.006

Norman, D. A., & Shallice, T. (1986). Attention to action: willed andautomatic control of behaviour. In R. J. et al. Davidson (Ed.), *Consciousness and Self-regulation(Advances in Research and Theory, Vol. 4)* (pp. 1–18).

Over, H., & Carpenter, M. (2013). The social side of imitation. *Child Development Perspectives*, *7*(1), 6–11.

Over, Harriet, & Carpenter, M. (2009). Priming third-party ostracism increases affiliative imitation in children. *Developmental Science*, *12*(3), F1–F8. https://doi.org/10.1111/j.1467-7687.2008.00820.x

Over, Harriet, & Carpenter, M. (2012). Putting the social into social learning: Explaining both selectivity and fidelity in children's copying behavior. *Journal of Comparative Psychology*, *126*(2), 182–192. https://doi.org/10.1037/a0024555

Over, Harriet, Carpenter, M., Spears, R., & Gattis, M. (2013). Children selectively trust individuals who have imitated them. *Social Development*. https://doi.org/10.1111/sode.12020

Pietikäinen, K. S. (2014). ELF couples and automatic code-switching. *Journal of English as a Lingua Franca*, *3*(1), 1–26.

Pinker, S. (2010). The cognitive niche: Coevolution of intelligence, sociality, and language. *Proceedings of the National Academy of Sciences*, *107*(Supplement_2), 8993–8999. https://doi.org/10.1073/pnas.0914630107

Rakoczy, H., Brosche, N., Warneken, F., & Tomasello, M. (2009). Young children's understanding of the context-relativity of normative rules in conventional games. *British Journal of Developmental Psychology*. https://doi.org/10.1348/026151008X337752

Rakoczy, H., Warneken, F., & Tomasello, M. (2008). The sources of normativity: Young children's awareness of the normative structure of games. *Developmental Psychology*, *44*(3), 875–881. https://doi.org/10.1037/0012-1649.44.3.875

Rakoczy, H., Warneken, F., & Tomasello, M. (2009). Young children's selective learning of rule games from reliable and unreliable models. *Cognitive Development*. https://doi.org/10.1016/j.cogdev.2008.07.004

1471 Richerson, P. J., & Boyd, R. (2005). *Not by Genes Alone: How Culture Transformed Human*
1472 *Evolution*. Chicago: University of Chicago Press.

1473 Sabbagh, M. A., & Henderson, A. M. E. (2007). *How an Appreciation of Conventionality*
1474 *Shapes Early Word Learning*. (115), 25–38. https://doi.org/10.1002/cad

1475 Schmidt, M. F., Butler, L. P., Heinz, J., & Tomasello, M. (2016). Young children see a single
1476 action and infer a social norm: Promiscuous normativity in 3-year-olds. *Psychological*
1477 *Science*, *27*(10), 1360–1370.

1478 Schmidt, M. F. H., Rakoczy, H., & Tomasello, M. (2012). Young children enforce social norms
1479 selectively depending on the violator's group affiliation. *Cognition*, *124*(3), 325–333.
1480 https://doi.org/10.1016/j.cognition.2012.06.004

1481 Singh, D., & Bronstad, P. M. (1997). Sex differences in the anatomical locations of human
1482 body scarification and tattooing as a function of pathogen prevalence. *Evolution and*
1483 *Human Behavior*, *18*(6), 403–416.

1484 Sosis, R., & Alcorta, C. (2003). Signaling, Solidarity, and the Sacred: The Evolution of
1485 Religious Behavior. *Evolutionary Anthropology*. https://doi.org/10.1002/evan.10120

1486 Taniguchi, Y., & Sanefuji, W. (2017). The boundaries of overimitation in preschool children:
1487 Effects of target and tool use on imitation of irrelevant actions. *Journal of Experimental*
1488 *Child Psychology*. https://doi.org/10.1016/j.jecp.2017.01.014

1489 Tennie, C., Call, J., & Tomasello, M. (2009). Ratcheting up the ratchet: on the evolution of
1490 cumulative culture. *Philosophical Transactions of the Royal Society B: Biological*
1491 *Sciences*, *364*(1528), 2405–2415. https://doi.org/10.1098/rstb.2009.0052

1492 Tomasello, M., Kruger, A. C., & Ratner, H. H. (1993). Cultural learning. *Behavioral and Brain*
1493 *Sciences*, *16*, 495–511.

1494 Tomasello, Michael, Carpenter, M., Call, J., Behne, T., & Moll, H. (2005). Understanding and
1495 sharing intentions: The origins of cultural cognition. *Behavioral and Brain Sciences*,
1496 *28*(5), 675–691. https://doi.org/10.1017/S0140525X05000129

1497 Tooby, J., & Cosmides, L. (2005). Conceptual foundations of evolutionary psychology. In D.
1498 M. Buss (Ed.), *The handbook of evolutionary psychology* (pp. 5–67). Hoboken, NJ, USA:
1499 Wiley.

1500 Uzgiris, I. C. (1981). Two Functions of Imitation During Infancy. *International Journal of*
1501 *Behavioral Development*, *4*(1), 1–12. https://doi.org/10.1177/016502548100400101

1502 Vaish, A., Missana, M., & Tomasello, M. (2011). Three-year-old children intervene in third-
1503 party moral transgressions. *British Journal of Developmental Psychology*, *29*(1), 124–
1504 130. https://doi.org/10.1348/026151010X532888

1505 Vredenburgh, C., Kushnir, T., & Casasola, M. (2015). Pedagogical cues encourage toddlers'
1506 transmission of recently demonstrated functions to unfamiliar adults. *Developmental*
1507 *Science*. https://doi.org/10.1111/desc.12233

1508 Watson-Jones, R. E., Legare, C. H., Whitehouse, H., & Clegg, J. M. (2014). Task-specific

effects of ostracism on imitative fidelity in early childhood. *Evolution and Human Behavior*. https://doi.org/10.1016/j.evolhumbehav.2014.01.004

Watson-Jones, R. E., Whitehouse, H., & Legare, C. H. (2016). In-Group Ostracism Increases High-Fidelity Imitation in Early Childhood. *Psychological Science*. https://doi.org/10.1177/0956797615607205

Whitehouse, H. (1995). *Inside the Cult: religious innovation and transmission in Papua New Guinea*. Oxford, UK: Oxford University Press.

Whitehouse, H. (2004). *Modes of Religiosity: a cognitive theory of religious transmission*. Walnut Creek, CA: AltaMira Press.

Whitehouse, H. (2012a). Human rites: Rituals bind us, in modern societies and prehistoric tribes alike. But can our loyalties stretch to all of humankind? *Aeon*.

Whitehouse, H. (2012b). Ritual, Cognition and Evolution. In R. Sun (Ed.), *Grounding the Social Sciences in the Cognitive Sciences*. Cambridge, MA: MIT Press.

Whitehouse, H. (2000). *Arguments and Icons: divergent modes of religiosity*. Oxford: Oxford University Press.

Whitehouse, Harvey. (2011). The Coexistence Problem in Psychology, Anthropology, and Evolutionary Theory. *Human Development*, *54*(3), 191–199. https://doi.org/10.1159/000329149

Whitehouse, Harvey. (2018). Dying for the group: Towards a general theory of extreme self-sacrifice. *Behavioral and Brain Sciences*, *41*, 1–64. https://doi.org/10.1017/S0140525X18000249

Whitehouse, Harvey, & Lanman, J. A. (2014). The Ties That Bind Us. *Current Anthropology*, *55*(6), 674–695. https://doi.org/10.1086/678698

Whiten, A. (2017). Social Learning and Culture in Child and Chimpanzee. *Annual Review of Psychology*, *68*(1), 129–154. https://doi.org/10.1146/annurev-psych-010416-044108

Whiten, A., Allan, G., Devlin, S., Kseib, N., Raw, N., & McGuigan, N. (2016). Social Learning in the Real-World: 'Over-Imitation' Occurs in Both Children and Adults Unaware of Participation in an Experiment and Independently of Social Interaction. *PLOS ONE*, *11*(7), e0159920. https://doi.org/10.1371/journal.pone.0159920

Whiten, A., Custance, D. M., Gomez, J.-C., Teixidor, P., & Bard, K. A. (1996). Imitative learning of artificial fruit processing in children (Homo sapiens) and chimpanzees (Pan troglodytes). *Journal of Comparative Psychology*, *110*(1), 3–14. https://doi.org/10.1037/0735-7036.110.1.3

Williamson, R. A., & Markman, E. M. (2006). Precision of imitation as a function of preschoolers' understanding of the goal of the demonstration. *Developmental Psychology*, *42*(4), 723–731. https://doi.org/10.1037/0012-1649.42.4.723

Wiltermuth, S. S., & Heath, C. (2009). Synchrony and cooperation. *Psychological Science*. https://doi.org/10.1111/j.1467-9280.2008.02253.x

Woodburn, J. (1982). Egalitarian Societies. *Man, NS*, *17*(3), 431–51.

Xygalatas, D., Mitkidis, P., Fischer, R., Reddish, P., Skewes, J., Geertz, A. W., … Bulbulia, J. (2013). Extreme Rituals Promote Prosociality. *Psychological Science*, *24*(8), 1602–1605. https://doi.org/10.1177/0956797612472910

Young, J. M., Krantz, P. J., McClannahan, L. E., & Poulson, C. L. (1994). Generalized Imitation and Response-Class Formation in Children with Autism. *Journal of Applied Behavior Analysis*. https://doi.org/10.1901/jaba.1994.27-685

Yu, Y., & Kushnir, T. (2014). Social context effects in 2-and 4-year-olds' selective versus faithful imitation. *Developmental Psychology*. https://doi.org/10.1037/a0034242